12-2008

Analysis of Speed-Related Crashes Using Event Data Recorder Data

Swathi Korpu
Clemson University, korpuswathi@gmail.com

ANALYSIS OF SPEED-RELATED CRASHES USING EVENT DATA RECORDER DATA

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Civil Engineering

---

by
Swathi Korpu
December 2008

Accepted by:
Dr. Jennifer H. Ogle, Committee Chair
Dr. Wayne Sarasua, Committee Member
Dr. Mashrur Chowdhury, Committee Member

ABSTRACT

Safety in general is defined as "the absence of unintended harm to living creatures or inanimate objects" (Evans, 2004). Traffic safety is measured in terms of motor vehicle crashes. A crash is an unplanned event which results in either fatality, or injury or property damage. Of all the factors that contribute to a crash, speed is a major issue, as more than 31 percent of all fatal crashes involve speeding (NHTSA, 2007). The speed-safety relationship can be determined by knowing the effect of speed on the frequency and severity of crashes. The relationship between speed and frequency of crashes is much more complicated than that between speed and severity of crashes which is based on principle of physics. Many researchers have studied of these relationships, but there is a risk of validation of the results so obtained, as the data used for most of these studies was obtained from police reports, self reports, case-control studies etc which may be erroneous. As a result, there is a need to revisit these relationships using more accurate and reliable data, and to determine the consistency between police reported speed and actual travel speed. If these two sources have a high level of correlation, there would be more trust placed in findings of older research. Event Data Recorders (EDR), installed in most current vehicles with airbags are capable of capturing and storing pre-crash, crash and post-crash information is proved to be one of the best sources available today.

This research focuses on analyzing the crashes for the years 2002 and 2003 using the supplemental EDR data from the National Automotive Sampling System / Crashworthiness Data System (NASS-NASS/CDS) database collected/maintained by the National Highway Traffic Safety Administration (NHTSA). All crashes in database were filtered to identify a set of records with complete EDR data. From this selection set, analysis was undertaken to include descriptive statistics and comparisons of speed vs type of crash, severity of crash, age and sex of driver, etc. It is observed from the descriptive statistics that EDR subset has some bias that is not apparent in overall NASS sample. It is also observed that the speeding crashes showed the same trend as earlier. For example, high involvement of young drivers, high amount of speeding in case of DUI, etc. Finally, it is found from the statistical tests that the speed five seconds prior to the crash is a better indicator of chosen speed, and that the difference in the average of police reported speed and speed at-5 seconds is practically significant.

DEDICATION

This thesis is dedicated to my parents, Narasimha Vara Prasad Korpu and Parvathi Korpu who inspired me and supported me in all aspects.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor Dr. Jennifer Harper Ogle for her support and encouragement throughout my master's study. Her patience and understanding nature inspired me from the beginning. With her continuous advice and direction, I successfully completed this thesis.

It is my pleasure to express my gratitude to my committee members Dr. Mashrur Chowdhury and Dr. Wayne Sarasua for their friendly advices and support.

I'm very much thankful to Chamani for her help, support and caring for me. She is one of my best friends and because of her I was able to withstand my homesickness.

I would like to thank my classmates Priyanka, Sukumar, Aimee, Tabreez, Vijay, Johnattan, Jo Ann and Pranoy for making me confident and helping me in various aspects.

I would like to thank my friends Srimathi, Gayatri, Sowjanya, Lahari and Anusha for their friendship and support.

I would also like thank my best friend Venkatesh Katragadda for his valuable help and company.

I would like to express a special gratitude to my family. To my father and mother, for their support, care and love for me. To my sisters and brother, for their love and friendliness. The love and support from my family is invaluable.

Finally, I would like to thank my loving friend Anil Kumar Puli who contributed his time and support for my success. I would not be able to complete this thesis without him.

TABLE OF CONTENTS

Page

Table of Contents (Continued)

Page

LIST OF FIGURES

Figure Page

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

LIST OF TABLES

Table | Page

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the United States, 90 percent of all transportation related fatalities are motor vehicle fatalities (NHTSA, 2008a). In the year 2006, 42,642 people were killed, 2,575,000 people were injured, and another 4,189,000 crashes involved property damage for an estimated total of 5,973,000 police-reported motor vehicle traffic crashes. On average, 117 people died each day (i.e., one every 12 minutes) in motor vehicle crashes in 2006 (NHTSA, 2008a).

Many different factors including driver, vehicle and environment factors can be contributing in whole or in part for motor vehicle crashes. Driver-related factors are mostly behavioral in nature and involve speeding, distracted driving, etc. (NHTSA, 2005). This study deals with the analysis of speed related traffic crashes in United States using the sample data collected by National Highway Traffic Safety Administration (NHTSA).

Speeding is one of the major contributing factors responsible in traffic crashes. A crash is considered to be speeding-related if the driver was charged for violating the posted speed limit or if an officer reports driving too fast for conditions (NHTSA, 2007). According to NHTSA's Traffic Safety Facts 2006, speeding is a contributing factor in 31 percent of all fatal crashes and the economic cost of speeding related crashes is estimated to be $40.4 billion each year.

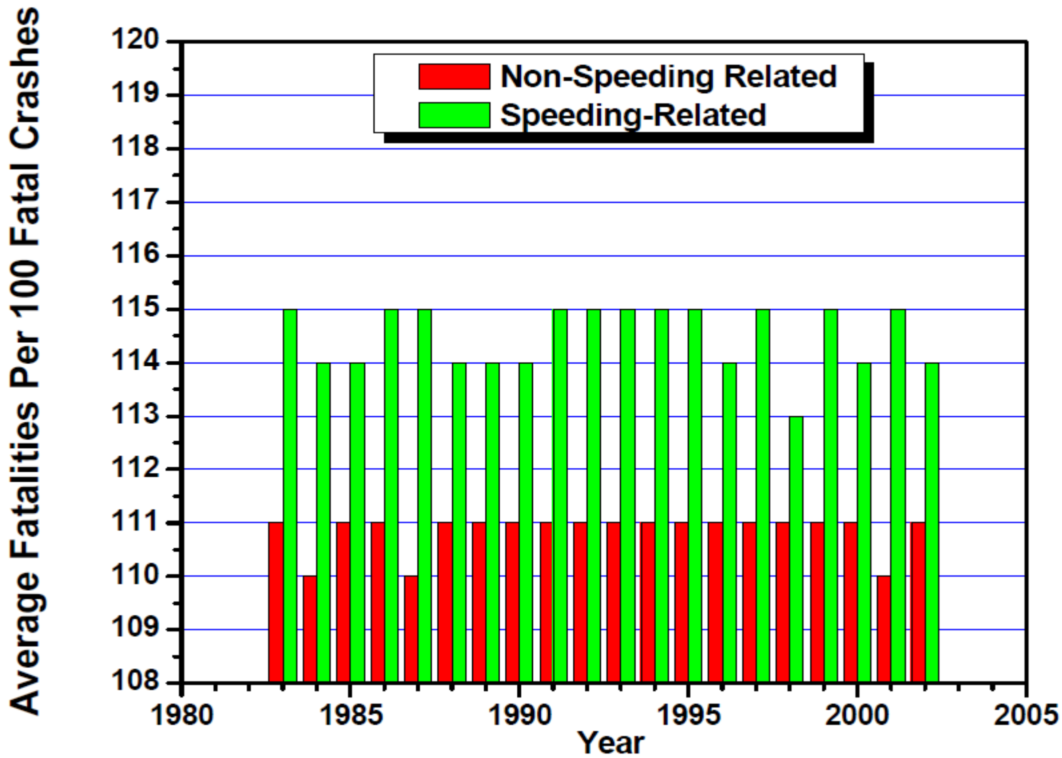


Figure 1.1: Figure showing the extent of speeding related fatalities (NHTSA, 2005)

Many factors influence the speeds drivers chose to travel. These factors can be grouped in to three categories namely driver attitudes and behavior (e.g. age and aggressiveness), road characteristics (e.g. sight distance) and environmental conditions (e.g. weather) (FHWA, 1998). In this report, the influence of some of these factors on speeding crashes is studied.

It should be noted that travel speed also depends on posted speed limit. Figure 1.2 shows the number of speeding related fatalities at various speed limit ranges. Much of the analysis in this study is done with the posted speed limit as a reference.

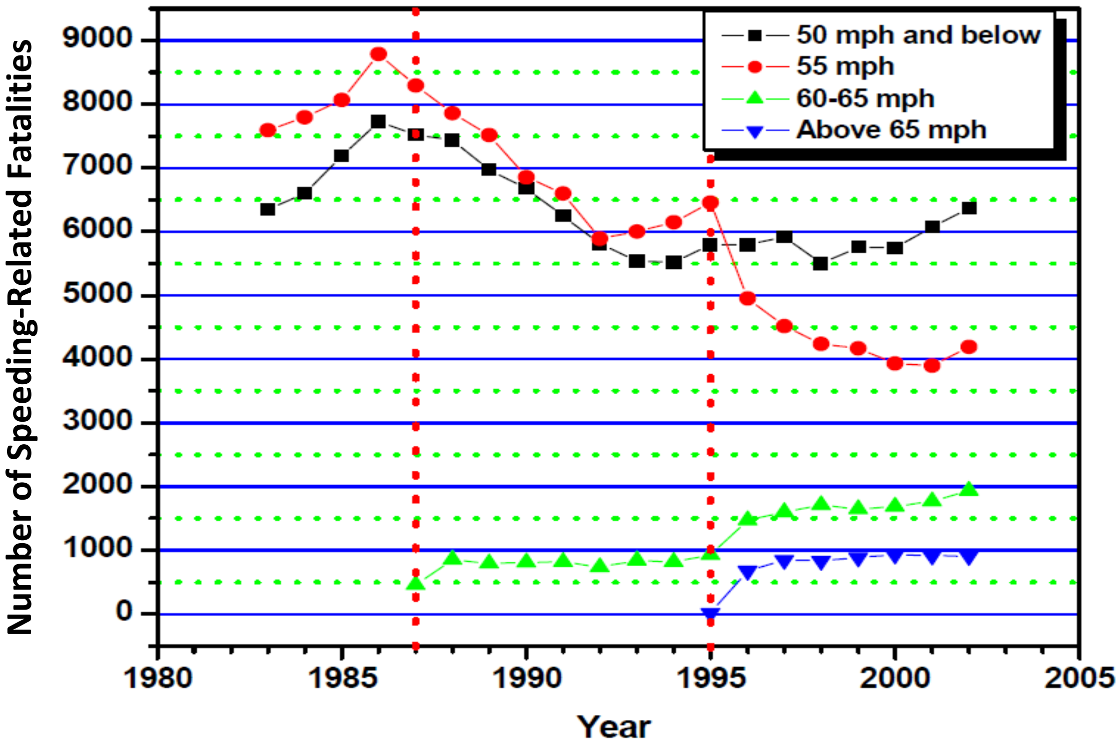


Figure 1.2: Speeding related fatalities by speed limit (NHTSA, 2005)

Figure 1.2 clearly explains the effect of posted speed limit on safety. In 1987, Congress permitted states to increase the posted speed limit to 65 mph and removed the National Maximum Speed Limit (NMSL) in 1995 (NHTSA, 2005). So after 1987, the number of speeding-related fatalities on roads with posted speed limits 65 mph remained the same; whereas there was a decrease of these numbers on roads with posted speed limits 55 mph and under. After the NMSL was removed, the number of fatalities has been increasing on roads with posted speed limits 65 mph and above. The fatalities have been same on roads with posted speed limits 50 mph and under. An abrupt drop of fatalities on roads with speed limit 55 mph is due to the miles of these roadways (NHTSA, 2005).

Many researchers have conducted studies (explained in Chapter 2) to understand the speed-safety relationship. The data used in these studies we typically obtained from police reports, case control methods and accident reconstruction techniques.

Unfortunately, these data may be somewhat subjective and prone to error, thus leading to the risk of validation of the results. To overcome these issues, there is a need for more accurate pre-crash speed data.

Event Data Recorders (EDR) installed in most current vehicles with airbags are able to record and store the pre-crash, crash and post-crash information. EDR are capable of recording/storing a number of variables for later analysis.

An EDR can record information (variable data) for two types of events namely deployment and non-deployment which are both associated with airbags deployment. A non-deployment event is an event that is "severe enough to 'wake up' the airbag sensing algorithm but not severe enough to deploy the air bag(s)" (NHTSA, 2008b). A deployment event on the other hand, is an event that is severe enough to deploy the airbags.

Current databases mainly contain data from GM vehicles and Ford vehicles because these manufacturers have released codes to read the data from engine computer where other manufacturers have not. NHTSA stores this information for three crash programs National Automotive Sampling System Crash Worthiness Data System (NASS-NASS/CDS), Special Crash Investigations (SCI) and Crash Injury Research and Engineering Network (CIREN). The pie chart shown below gives the fraction of EDR files distributed per each of these crash programs.

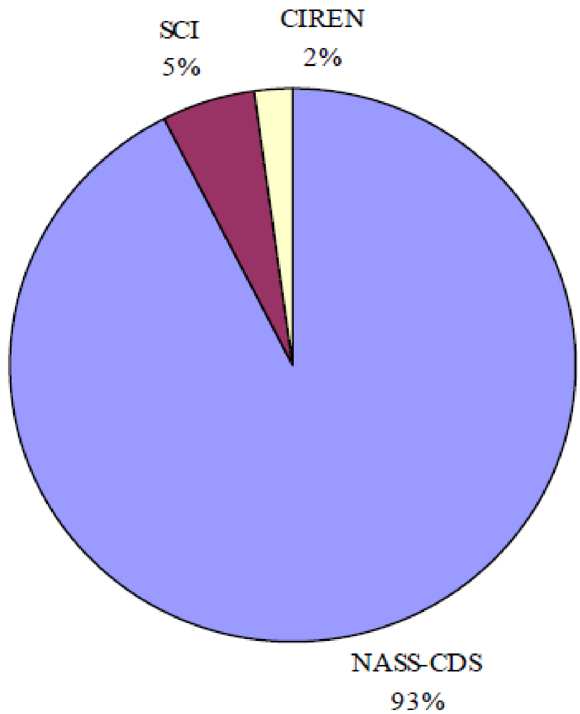


Figure 1.3: EDR file distribution per NHTSA crash program (NHTSA, 2008b)

EDR data in combination with standard crash reporting elements make a power set of data for analysis of speed/crash relationships. This data can be categorized into two groups namely Quantitative and Qualitative which deals with numbers and descriptions respectively.

1.2 Problem Statement

More than 42,000 people die on U.S. Highways each year. Speeding contributed to more than 31% of all fatal crashes (NHTSA, 2007). A number of researchers have studied the speed-safety relationship, using estimated pre-crash speed obtained from police reports, witnesses and crash investigation.

Until recently, our ability to obtain accurate and reliable pre-crash speed information has been hampered by technology/cost. Currently, most vehicles are equipped with EDRs which are capable of collecting a set of vehicle operating parameters just prior to and during a crash event. EDRs can record and store the speed of the vehicle five seconds prior to crash in one second intervals and as well, a delta velocity change for the milliseconds prior to/during crash. The accuracy and resolution of the data obtained from EDR is proven to be acceptable (See Chapter2). Further, researchers have used the Delta-V (difference in velocity before and after impact) obtained from the EDRs to study crash severity and injury correlations with success.

1.3 Objectives

The problem being described, the following are the objectives of this study:

1. To determine if police reported speeds are accurate; and if not, whether a correlation exists.
2. To examine the relationship between speed and various factors (age and sex of the driver, speed limit, roadway characteristics, alcohol etc.) that have historically been associated with speed.

1.4 Organization of the Thesis

Chapter two provides a two part Literature review. The first part of literature review gives a brief description of previous studies on finding the relationship between speed and safety; and the limitations associated with those studies. The second part of literature review describes the availability of Event Data Recorder data---its advantages and disadvantages. Chapter 3 Methodology explains the data source and the format of the data that is used, the process of querying and the statistical tests that are done. Chapter 4 provides the results obtained from the analysis of data. Analysis is done in terms of the variables (factors) that are supposed to influence the speed. The results are shown in the tabular and graphical formats as well as SAS outputs. Chapter 5 gives the summary of conclusions (findings) that provides the causes and effects of speeding and thereby providing scope of present scenario and some future recommendations to improve traffic safety.

CHAPTER 2

LITERATURE REVIEW

This chapter deals with the influence of speed on crashes and their severity and explains the need to collect the onboard crash data to improve traffic safety. The Literature Review for this thesis is divided into two parts.

- Part I explains the effect of speed on safety and
- Part II provides necessary background on Event Data Recorder (EDR) and accuracy.

According to Evans (2004), “A vehicle striking anything is referred to as a crash.” Crashes have traditionally been used as a measure of safety. In other words, fewer the number of crashes, better the level of safety. Thus, the main source of traffic safety information is the crash data.

Speed is one of the most prevalent contributing factors in motor vehicle crashes. The following facts were taken from NHTSA’s Safety Facts 2006.

- Speeding is associated with approximately 31 percent of all fatal crashes.
- Each year, the estimated cost of speeding-related crashes is $40.4 billion.
- Of all fatal crashes involving 15-to-20-year-old male drivers, 39 percent are due to speeding.
- 76 percent of speeding drivers involved in fatal crashes are drunk between midnight and 3 a.m.

- Speeding related fatalities on Interstates account for only 13 percent of total fatalities.

To fully comprehend previous research, an understanding of different terminology is very important. The definitions of various speed terms are provided in the following table:

Table 2.1: Definitions of various speeds

Measure	**Definition [**MUTCD (2000,as cited in Ogle,2005)]
Design Speed	**The Design Speed** is a selected speed used to determine the various geometric design features of the roadway.
85th Percentile Speed	**85th Percentile Speed** is the speed at or below which 85 percent of the motorized vehicles travel.
Operating Speed	**Operating Speed** is the speed at which a typical vehicle or the overall traffic operates. Operating speed may be defined with speed values such as the average, pace, or 85th percentile speeds.
Average Speed	**Average Speed** is the summation of the instantaneous or spot-measured speeds at a specific location of vehicles divided by the number of vehicles observed.
Pace Speed	**Pace Speed** is the highest speed within a specific range of speeds that represents more vehicles than in any other like range of speed. The range of speeds typically used is 10 km/hr or 10 mph.

Posted speed limit is highly correlated with operating speeds and thereby related to safety. According to NCHRP Report 504 (Fitzpatrick et al., 2003), for a given section of highway, the posted speed limit is the pertinent maximum or minimum speed determined by law. In general, 85th percentile operating speed is used to set the posted speed limit. The 85th percentile speed is influenced by various design elements, roadside and roadway variables.

Fitzpatrick et al. (2003) collected speeds of free-flowing vehicles during dry pavement conditions on week days between 7:00 am and 6:00 pm. For all functional classes of roadways and speed ranges, they found that a strong relationship exists between operating speed and posted speed limit. A model was developed for 85th percentile speeds Q_{85}

$$E[Q_{85}] = 7.675 + 0.98 \times \text{Posted Speed Limit}$$

where $E[Q_{85}]$ is the expected value of Q_{85}. The following are acknowledged from this linear equation.

- The 85th percentile speed varies at the same rate with posted speed limit (as the slope 0.98 is very close to one).
- The 85th percentile speed is higher than the posted speed limit approximately by 7 to 8 mph (as the constant is 7.675).

Further they found that the models--- Q_{95}, Q_{90}, Q_{85}, Q_{50}, and Q_{15} are also highly correlated with the posted speed limit.

2.1 Effect of Speed on safety

According to Federal Highway Administration (FHWA, 1998) vehicle speed is related to traffic safety in the following two ways.

1. Speed and Occurrence of crashes: Increase in vehicle speed results in the increase of the driver's perception-reaction time which in turn leads to the increase of probability for a crash to happen.
2. Speed and Severity of crashes: Increase in the vehicle speed results in the increase of kinetic energy which is a function of mass and vehicle speed.

2.1.1 Speed and occurrence of crashes

The relationship between speed and crash involvement risk is complicated when compared to the relationship between speed and crash severity. The following studies are some of the most referenced works on speed-crash risk relationship.

Solomon (1964) in a study of speed and crashes on two-lane and four-lane highways compared the speeds of crash involved vehicles with the speeds of non-crash involved vehicles. He found that the crash rates increases with increase in deviation of travel speeds from the mean speed of the traffic. Also, the results showed that the risk of being involved in a crash for drivers travelling at low speeds is high when compared to those travelling at relatively high speeds. He determined a U-shaped relationship between speed and crash rate which shows that not only speed but also the speed deviation increases the risk of involving in a crash.

Cirillo (1968) confirmed Solomon's results limiting his analysis to daytime crashes involving two or more vehicles travelling in the same direction on rural and urban Interstate freeways. He also found that the crash involvement rates at all speeds were lower when compared to those on two-way two-lane roads and thereby concluded that the roadway geometry also has a significant effect on crash probability.

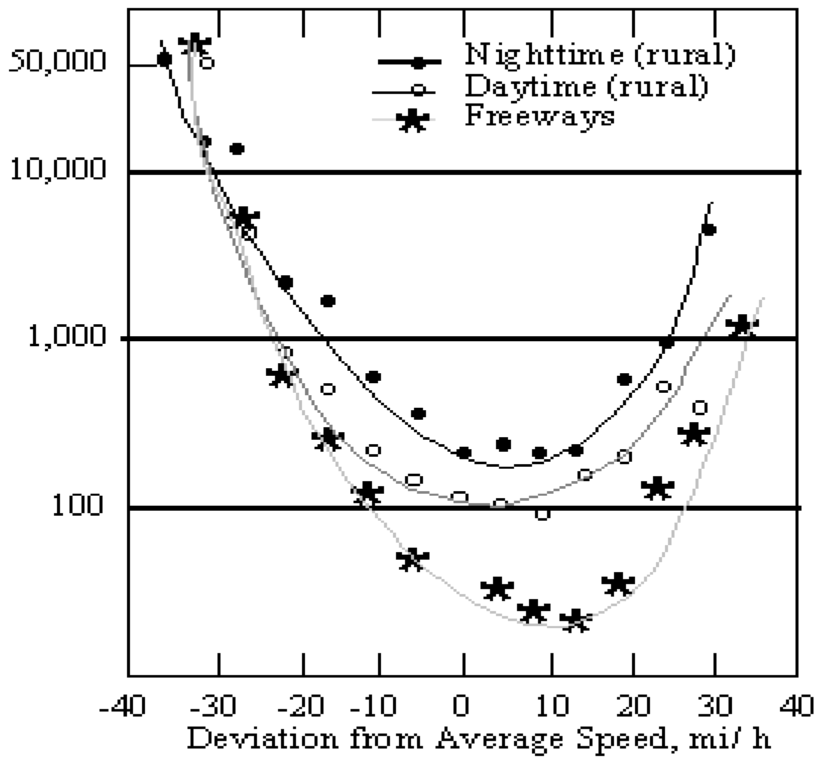


Figure 2.1: Crash involvement rate by deviation from average travel speed by Solomon (FHWA, 1998)

The results of these studies showed that the crash risk increases with increase in deviation from the mean speed. In other words, for vehicles traveling with speeds much higher and much lower than the mean speed are more likely to be involved in a crash. In these studies, the pre-crash speeds were estimated using police reports, witness reports or crash reconstruction techniques. Each of these methods have limitations and may not be accurate thereby leading to erroneous results. Some researchers used case control methods in which the speed data is collected in the conditions reflecting the crash situation (location, day of week etc.,) which may sometimes fail to match conditions of the actual time of the crash. In case control methods, the speed data that is collected does not include the vehicles that are turning or stopped. Moreover, the data was collected excluding intersections, driveways etc., that have a major effect on speed and unfortunately many of the slower vehicles involved in crashes may be stopped or slowed down for a maneuver.

To overcome these issues, Research Triangle Institute (RTI) in 1970 using the estimated pre-crash speeds from the police reports and also the speed data collected by means of the sensors embedded in the pavement to obtain the speed at the time of the crash investigated 114 crashes involving 216 vehicles on state road and found the less pronounced U-shaped relationship. West and Dunn (1971) reported the results of RTI. They compared the speeds of crash-involved vehicles with speeds of non-crash-involved vehicles and found that U-shaped relationship was greatly attenuated when the crashes involving turning vehicles were excluded from the sample. They also found that crash risk for vehicles traveling two standard deviations more than the mean speed is high.

Hauer (1971) questioned the results of the previous research mentioned above, by relating crash involvement with overtaking. He did not directly relate the crash risk and deviation from the mean speed. That is, he found that the rate of crash involvement increases with overtaking rate which increases with increase in deviation from the mean speed. In other words, he explained that there was not a real root cause identified in previous research because, he could get same results assuming an overtaking model.

Trying to get around some of the problems/issues of the previous research, Harkey et al. (1990) in their study of rural and urban roads with posted speed limits between 25 and 55 mph in North Carolina and Colorado considered only the weekday, non-alcohol, non-intersection crashes. They also found the same U-shaped curve as earlier; and all the above findings are represented in the figure below.

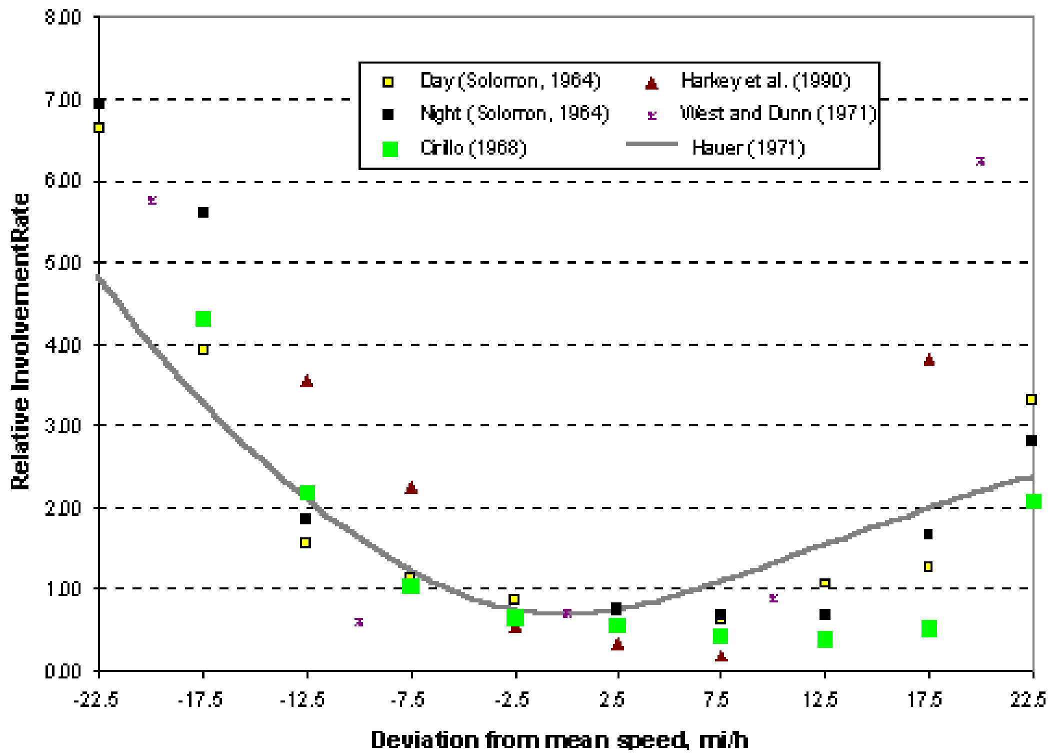


Figure 2.2: Crash involvement and overtaking rates relative to average rate and speed (FHWA, 1998)

The effect of road characteristics on speed is emphasized by Garber and Gadiraju (1988). They concluded that for all classes of roadways, the increase in speed variance results in increase of crash rate and that the speed variance is significantly influenced by the difference in design speed and posted speed limit and it is minimum when this difference is between 5 and 10 mph. They also reported that on roads with higher design speeds, the drivers tend to speed irrespective of the posted speed limit.

2.1.2 Speed and severity of crashes

A vehicle undergoes a rapid change in speed when it crashes. Vehicle occupants continue to move at the vehicle's previous speed until they are stopped either by striking the interior of the vehicle, by impact with objects external to the vehicle if ejected, or by being restrained (Evans, 1991). The probability that an occupant is injured increases with increase in the kinetic energy released during impact. Therefore, delta velocity (Delta-V) known as change in speed at impact is an important quantity in determining crash severity. The following are some of the studies that explain the speed-crash severity relationship.

Solomon (1964) reported that the higher the speed, the greater is the severity of the crash. For example, up to about 45 mph (72 km/h), 20 to 30 persons were injured and about 1 person killed per 100 crash-involved vehicles whereas, at speeds of 73 mph (117 km/h) and greater, nearly 130 persons were injured and 22 persons were killed per 100 crash-involved vehicles (Solomon 1964).

Using the data from the National Crash Severity Study O'Day and Flora (1982) investigated about 10,000 crashes and found that with the increase in change in speed (ΔV), the probability of fatality increases. Similar relationship was found from the Joksch (1993) study in which the data from the NHTSA's National Analysis Sampling System (NASS) is used. He found that the probability of a driver being involved in a fatal crash increases with increase in Delta-V to the fourth power. The figure shows the fatality risk curves developed from these two studies.

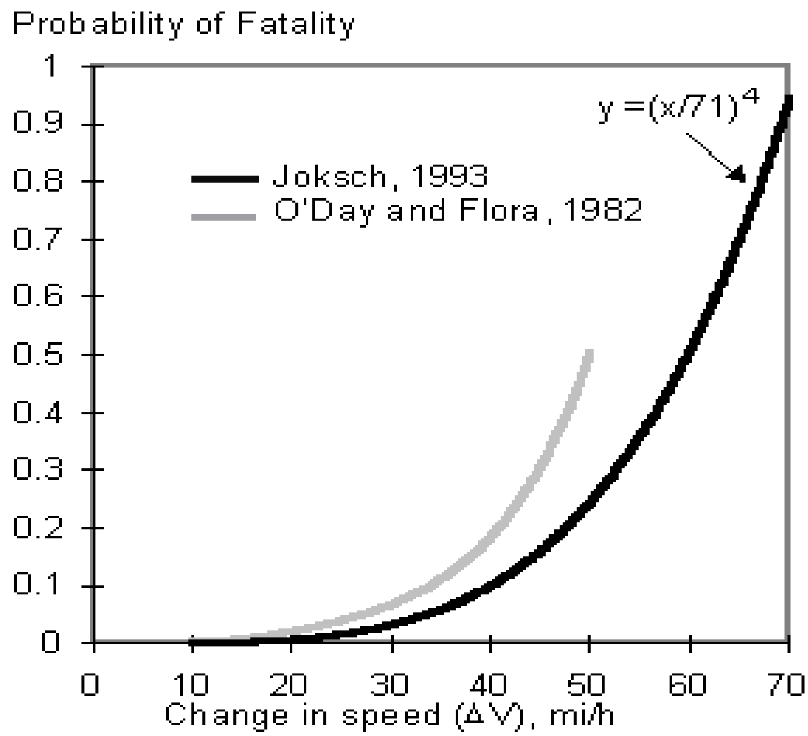


Figure 2.3: Effect of Delta-V on fatality risk (FHWA, 1998)

Bowie and Waltz (1994) using the NASS data from 1982-1989 and the Abbreviated Injury Scale (AIS) found the relationship between Delta-V and the crash severity. The tabulated result of their study which is shown below clearly supports the statement that the crash severity or injury levels increases with increase in delta velocity at the impact.

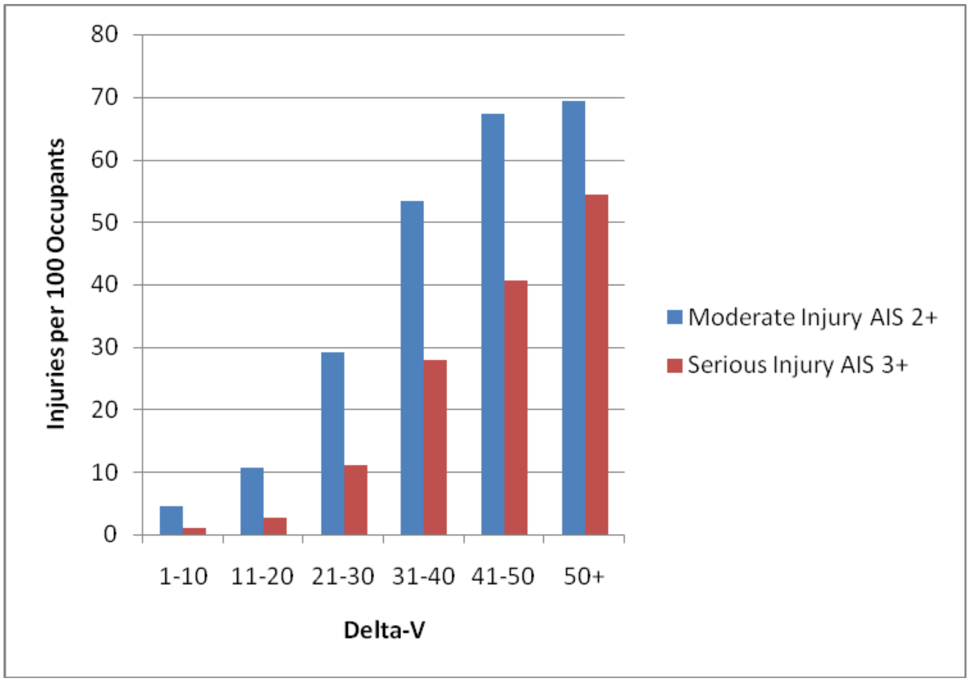


Figure 2.4: Graphical representation of Delta-V effect on injury severity (FHWA, 1998)

Another research was conducted in Australia by Kloeden et al. (1997) using the speed zones with 37mph speed limit. They compared the pre-crash speeds found using the computer-aided crash reconstruction technique with the free speeds measured under the case control environment and found that the probability of being involved in an injury crash increases exponentially at speeds higher than the median speed (37mph or 60km/hr) whereas it is lower for speeds at or below the median speed.

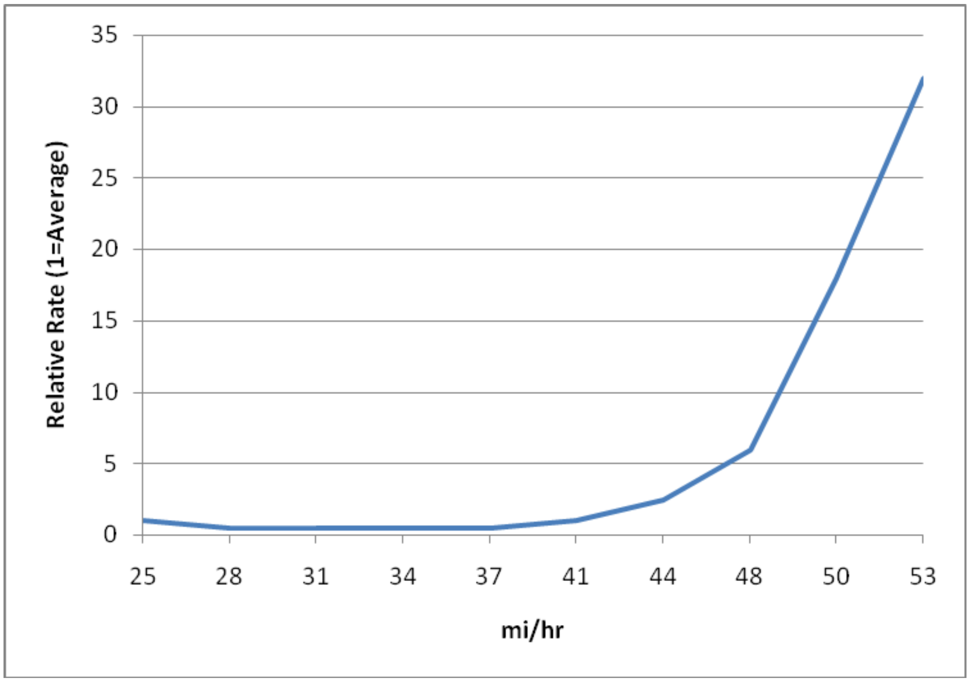


Figure 2.5: Injury crash involvement rate relative to average rate and travel speed (FHWA, 1998)

Previous research on Speed-Safety relationship described so far used the speed and crash data obtained from the police reports, driver reports, case-control studies, crash reconstruction techniques etc., which may/may not be accurate. In order to overcome the risk of validation of the results, accurate and reliable data is to be used for the analysis, which further helps in improving the road safety. Event Data Recorder (EDR) installed in a vehicle to control airbag functions is one of the best sources of pre-crash speed. EDRs have the capability to capture and store both the pre-crash, crash and post-crash data, which could be used for many of the highway crash data analysis done by the researchers, to enhance driver, vehicle and road side safety.

2.2 Event Data Recorder (EDR)

General Motors (GM) introduced Event Data Recorders for the first time in selected models in 1974, which were mainly used to control and record the deployment of air bags (NHTSA, 2008b). Since then, several features have been added to the EDRs, which now, are capable of recording and storing the crash relevant data---vehicle speed, engine speed, throttle position, and brake switch circuit status (ON/OFF), seat belt at one-second intervals up to five seconds prior to the crash.

In general, EDRs consist of four main components --- a sensory package, associated processors, storage of generated data, and a retrieval mechanism (NHTSA, 2008b).

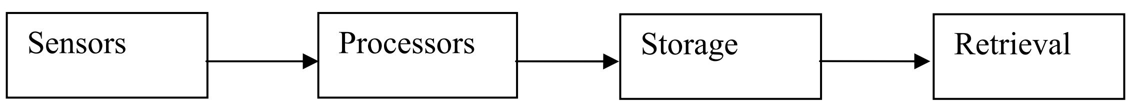


Figure 2.6: Components of EDR (NHTSA, 2008b)

If the Sensing and Diagnostic Module (SDM- the EDR in General Motors' vehicle), recognizes a potential crash, it monitors the acceleration-time history of the vehicle, and based on a built-in algorithm decides when air bags should be deployed (German, n.d.). Certain collision data is stored in the available computer memory, and is accessed using the Crash Data Retrieval (CDR) system. This system consists of interface cables that allow the user to connect directly between a laptop and SDM. The data retrieved is in hexadecimal format. The software in the CDR system interprets the data and displays it in tabular and graphical format (German, n.d.).

Table 2.2: Table showing collision data retrieved from EDR (NHTSA, n.d.)

SIR Warning Lamp Status	OFF
Driver's Belt Switch Circuit Status	BUCKLED
Passenger Front Air Bag Suppression Switch Circuit Status	Air Bag Not Suppressed
Ignition Cycles At Deployment	2836
Ignition Cycles At Investigation	2842
Maximum SDM Recorded Velocity Change (MPH)	-6.58
Algorithm Enable to Maximum SDM Recorded Velocity Change (msec)	127.5
Time Between Non-Deployment And Deployment Events (sec)	N/A
Time From Algorithm Enable to Deployment Command Criteria Met (msec)	27.5

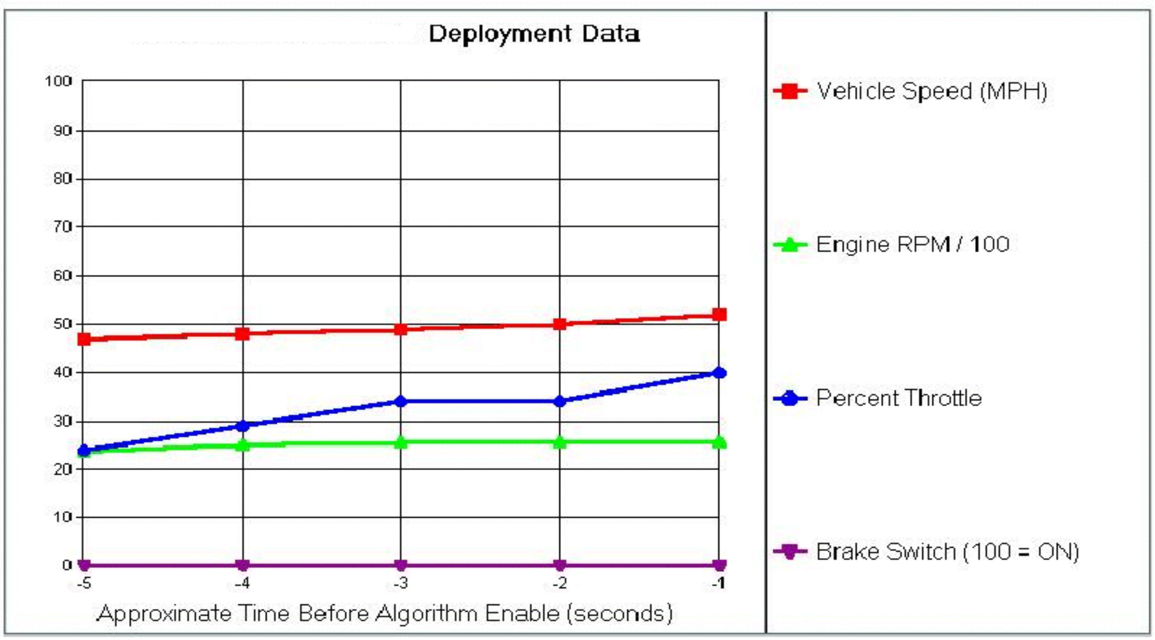


Figure 2.7: Graph showing pre-crash data retrieved from EDR (Chidester et al., 2001)

The CDR reports thus obtained are collected and stored by National Highway Traffic Safety Administration (NHTSA) as part of an effort to make the crash data bases more useful for the researchers to effectively analyze the motor vehicle crashes. National Automotive Sampling System / Crashworthiness Data System (NASS-NASS/CDS) is one of the NHTSA's crash data bases to which the EDR data is appended beginning in the year 2000. "NASS-NASS/CDS database is a national sample of 4,000 to 5,000 crashes investigated each year by NHTSA at 27 locations throughout the United States" (Gabler et al, 2004).

The figures below explain the functions associated with the above mentioned components of EDR. An example of a CDR report is shown in Appendix.

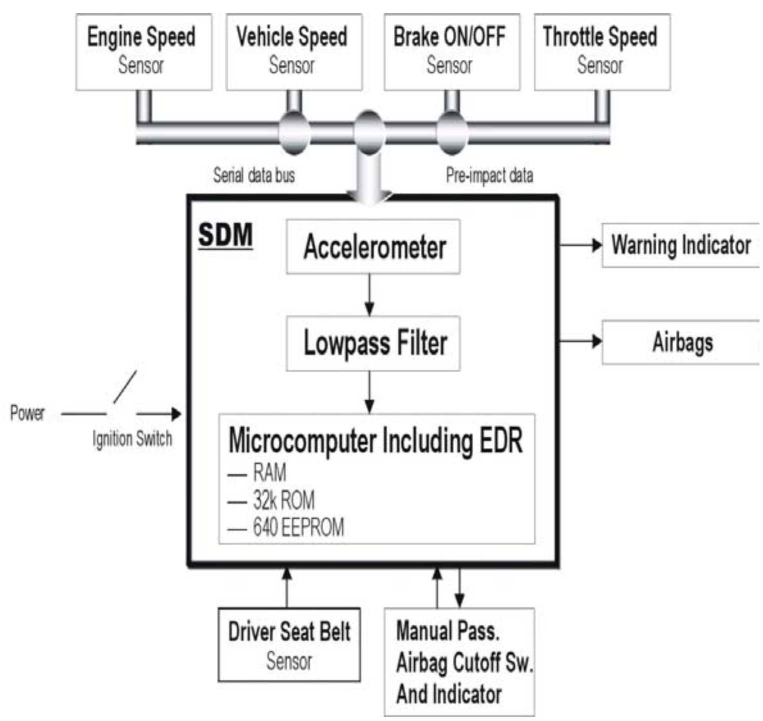


Figure 2.8: 1999 GM SDM Block Diagram (Chidester et al., 2001)

Figure 2.9: Data Retrieval from EDR (German, n.d.)

2.2.1 EDR Benefits, Accuracy and Limitations

To use the EDR data for research, it is also important to know the benefits and accuracy of the EDR data along with some of the associated limitations.

The benefits of EDR (or any on-board recorder) can be explained clearly using the Haddon matrix in which, the crash is divided into three segments namely pre-crash, crash and post-crash; and at the same time looks at the human, vehicle and environmental considerations for each segment. The following Haddon matrix shows the data that can be collected without EDR:

Table 2.3: Haddon matrix without EDR capability (Chidester et al., 1999)

Event/Factor	Human	Vehicle	Environment
Pre-Crash		Skid marks	
Crash		Estimated Delta-V	
Post-Crash	Injury	Collision damage	Environment after collision

In above matrix without EDR, no practical information is known for the vehicle, and what post crash and crash information that is avail is estimated given our best guess of conditions at time of the crash. With the use of EDR's numerous data can be collected. Haddon matrix shown below gives the potential pre-crash, crash and post-crash data that can be collected using an EDR.

Table 2.4: Haddon Matrix showing data elements collected by EDR (Chidester et al., 1999)

Event/Factor	Human	Vehicle	Environment
Pre-Crash	Belt use Steering Braking Occupant position Occupant weight	Speed Engine speed Percent throttle Brake application Anti-lock braking systems (ABS) Tire pressure Cruise control Other controls	Conditions during crash
Crash	Airbag data Pretensioners	Crash pulse Measured change in velocity Yaw Airbag activation time Roll angle	Location
Post-Crash	ACN(automatic collision notification)	ACN	ACN

EDR is found to be a source of accurate and reliable data. The table below shows the accuracy and resolution for the data recorded for the 1999 SDM. It can be observed that, the accuracy of vehicle speed recorded is ± 4% and the full scale (the maximum value that can be recorded) reading associated with it is 158.4 mph. The accuracies of Delta-V, engine speed, throttle position are also given in the table and are considerably accurate.

Table 2.5: Accuracy and Resolution of Data Recorded (Chidester et al, 1999)

Parameter	Full Scale	Resolution	Accuracy	How Measured	When Updated
ΔV	±55.9 mph	0.4 mph	± 10 %	Integrated acceleration	Recorded every 10msec, calculated every 1.25 msec.
Vehicle Speed	158.4 mph	0.6 mph	± 4%	Magnetic pickup	Vehicle speed changes by ≥ 0.1 mph.
Engine Speed	16383 RPM	¼ RPM	± 1 RPM	Magnetic pickup	RPM changes by ≥32 RPM.
Throttle Position	100% wide open throttle	0.4%	± 5%	Rotary potentiometer	Throttle position changes by ≥5%.

Previous research by Gabler et al. in 2004 identified some limitations associated with the EDRs, which are described as follows:

Limited Recording Times: Most of the EDRs have limited recording times available, to capture the entire event, especially in case of vehicle-to-roadside event. Therefore, EDRs must be capable of recording a minimum of 300 milliseconds of crash pulse, as well, pre-crash and post-crash data are also needed for the roadway safety analysis (Gabler et al, 2004).

Issues with multiple events: Certain crashes may have multiple events involved with them. Examination of number of events per vehicle for NASS 2000-2002 cases, showed that, crashes involving more than two events are around 46% and those involving three or more events are around 17% (Gabler et al, 2004). But, most of the EDRs are not capable of recording all the events and, for example, GM EDRs can record up to only two events. So the first of three events may be overwritten.

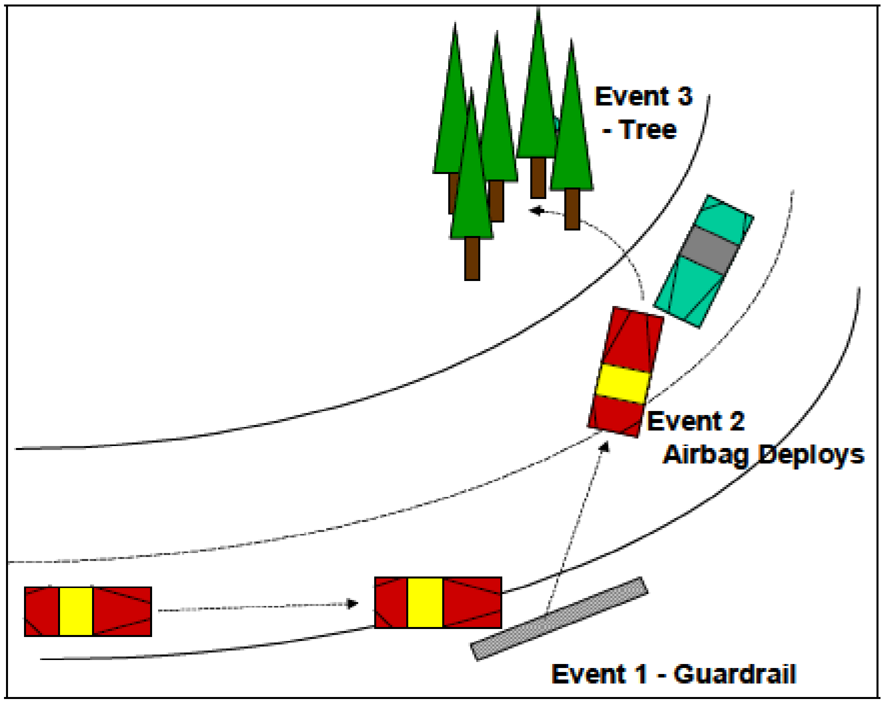


Figure 2.10: EDR's incapability in capturing multiple events (Gabler et al., 2004)

The problem of matching the EDR event with post-crash study: As mentioned earlier, a crash may involve multiple events and, as the EDR can record only some of them, it is important to know which events are captured by the EDR. In other words, as the recording is done at discrete intervals (-5,-4.-3,-2,-1), we obtain only discrete events. But, as braking can begin anywhere within one second period, it is necessary to have continuous data, which is not possible with EDR.

In summary, much research has been conducted on the speed crash relationship problem. The frequency of crashes involving speed alone indicates some relative involvement of speed with crashes. However, most speed data collected previously have numerous limitations which make the results of that research questionable. As a result, there is a need to find a source from which reliable and accurate data will be available. EDRs have the capability to record and store pre-crash, crash and post-crash data that could prove to be very useful in Highway Safety analysis. Furthermore, it is clear that EDRs provide accurate data in spite their limitations. So, accurate analysis and valid results can be achieved by using the EDR data in analyzing the speed related crashes.

CHAPTER 3

METHODOLOGY

The methodology for this research is carried out in different phases. The various phases are shown in the flow chart provided below.

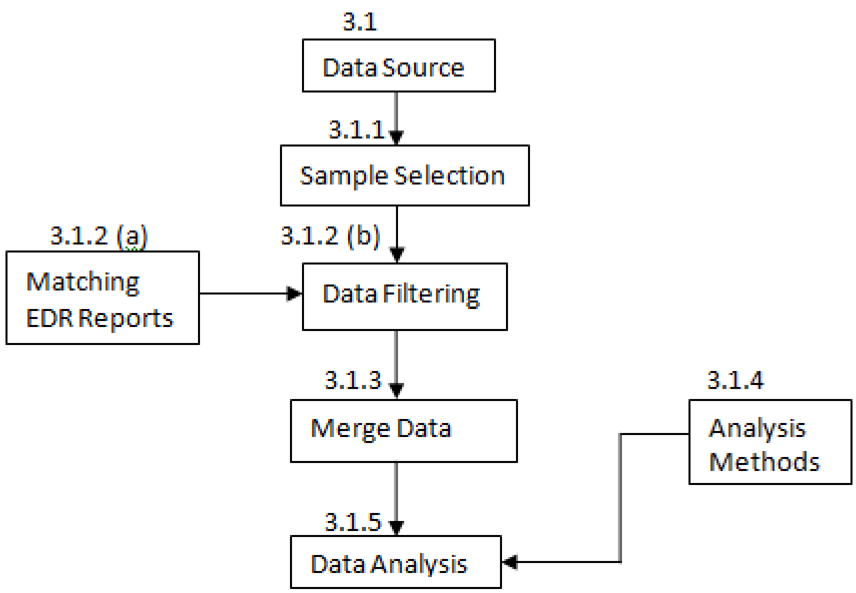


Figure 3.1: Flow chart showing various phases of methodology

<u>3.1 Data Source</u>

This thesis utilizes data from the National Automotive Sampling System/Crash-worthiness Data System (NASS-NASS/CDS) database collected by the National Highway Traffic Safety Administration. NASS-NASS/CDS is one of the NHTSA's vehicle crash investigation programs and contains a sample of 4,000 to 5,000 crashes per year investigated at 27 locations in United States by NHTSA (Gabler et al., 2004). "A NASS/CDS crash selected must be police reported, must involve a harmful event resulting from a crash and must involve at least one towed passenger car or light truck or van in transport on a traffic way" (NHTSA, 1998). The sampling is carried out in three stages namely selection of PSU's (PSU's represent the geographic areas of the entire country and are defined such that their minimum population was approximately 50,000), selection of police jurisdictions (a PSU contains a number of police jurisdictions which report the crashes that occur in that PSU and are selected based on measure of size) and, selection of crashes (the crashes in the jurisdictions are selected if they meet the requirement of NASS/CDS sampling) (NHTSA, 1998). The database consists of directories of crash data for each year (1997-2007) and a directory for EDR reports (2000-2006).

This research is conducted using data for the years 2002 and 2003. For these years, the EDR data from the EDR reports is transcribed manually [i.e. the data in the EDR reports (pdf files) is written into tables] and placed in a file named 'Expanded SAS data'. This task was completed by NHTSA to allow ease of use for safety research. The EDR data for these two years is available in a format that is compatible with the crash data. This file consists of tables that have the data elements that describe the crashes along with associated EDR information.

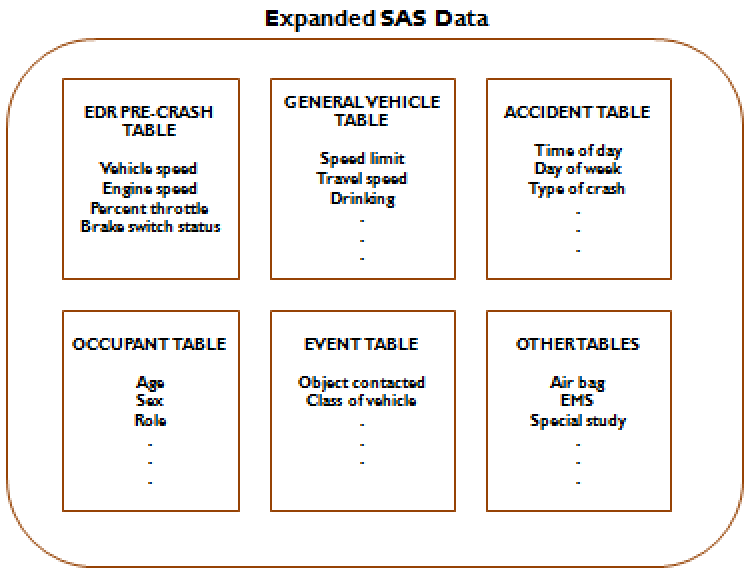


Figure 3.2: Tables and data present in expanded SAS data (NHTSA, 2002)

All the data tables available from NHTSA's NASS-NASS/CDS database are in SAS format, with 'sas7bdat' extension (e.g. accident.sas7bdat represents accident table), and is accessed using software named 'SAS System Viewer'. All these tables are having data view and variables view. Data view helps viewing the required information/values. Variable view provides the descriptions (including type, length, etc.) of all the variables indicated in data view. This software is useful only for viewing the data but not for any sort of data analysis. But, it has a feature that allows conversion of the data into CSV format which further can be imported into Microsoft Excel, MS Access, or other database format.

A snapshot of one of the tables viewed in this software is shown below. Arrow indicates the corresponding icon used to represent both views.

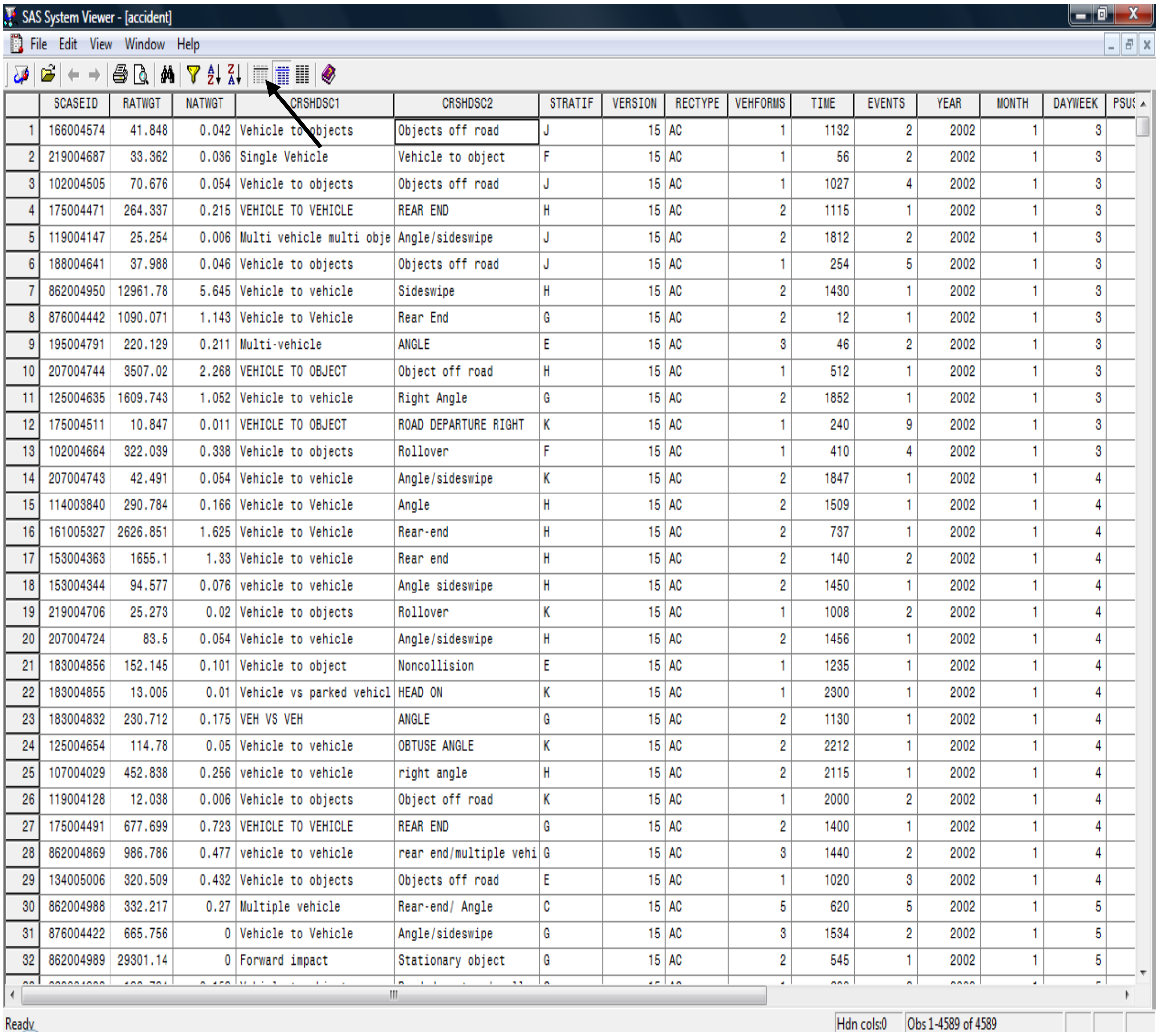
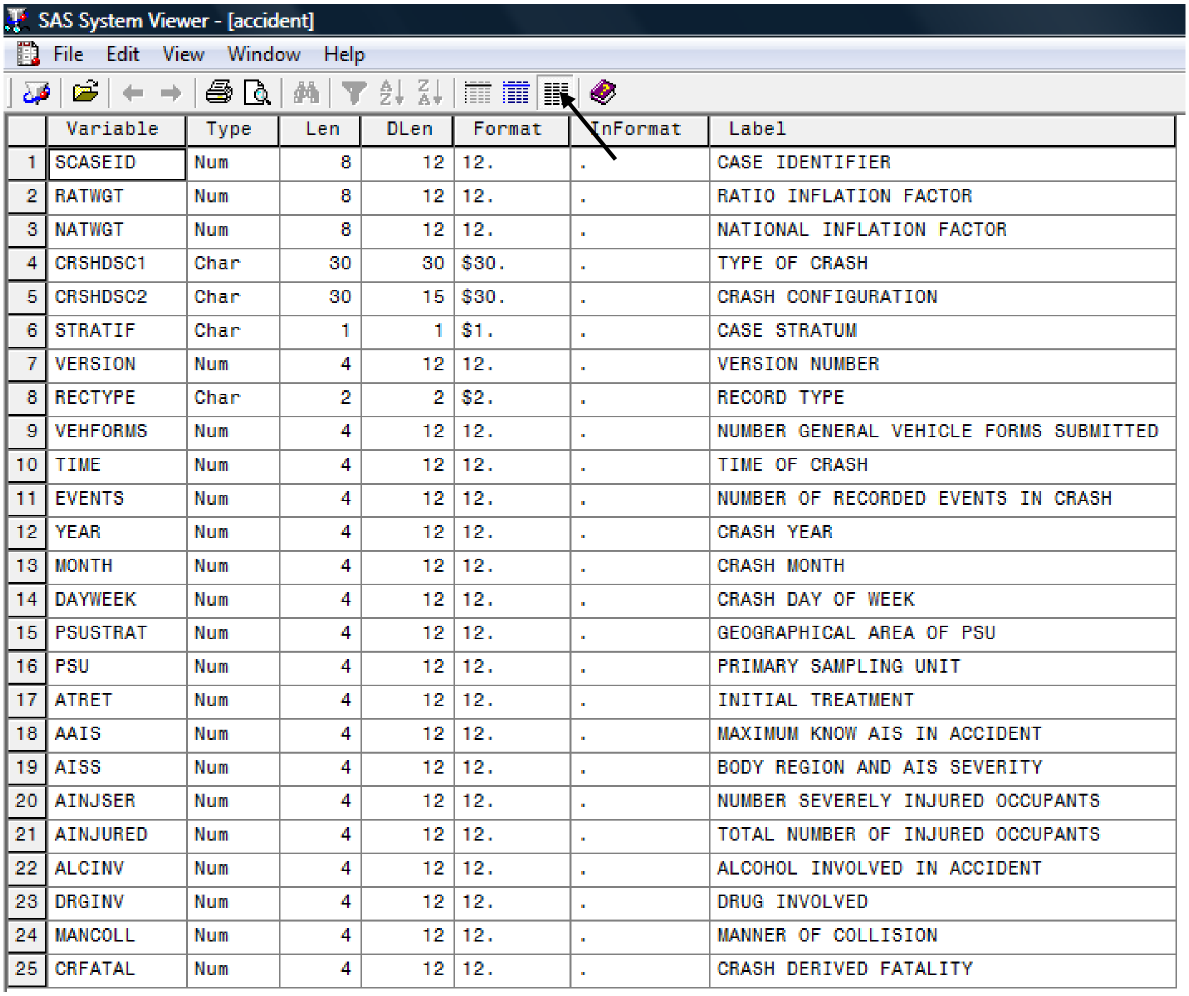


Figure 3.3: A data view of accident table in SAS System Viewer (NHTSA, 2002)



	Variable	Type	Len	DLen	Format	InFormat	Label
1	SCASEID	Num	8	12	12.	.	CASE IDENTIFIER
2	RATWGT	Num	8	12	12.	.	RATIO INFLATION FACTOR
3	NATWGT	Num	8	12	12.	.	NATIONAL INFLATION FACTOR
4	CRSHDSC1	Char	30	30	$30.	.	TYPE OF CRASH
5	CRSHDSC2	Char	30	15	$30.	.	CRASH CONFIGURATION
6	STRATIF	Char	1	1	$1.	.	CASE STRATUM
7	VERSION	Num	4	12	12.	.	VERSION NUMBER
8	RECTYPE	Char	2	2	$2.	.	RECORD TYPE
9	VEHFORMS	Num	4	12	12.	.	NUMBER GENERAL VEHICLE FORMS SUBMITTED
10	TIME	Num	4	12	12.	.	TIME OF CRASH
11	EVENTS	Num	4	12	12.	.	NUMBER OF RECORDED EVENTS IN CRASH
12	YEAR	Num	4	12	12.	.	CRASH YEAR
13	MONTH	Num	4	12	12.	.	CRASH MONTH
14	DAYWEEK	Num	4	12	12.	.	CRASH DAY OF WEEK
15	PSUSTRAT	Num	4	12	12.	.	GEOGRAPHICAL AREA OF PSU
16	PSU	Num	4	12	12.	.	PRIMARY SAMPLING UNIT
17	ATRET	Num	4	12	12.	.	INITIAL TREATMENT
18	AAIS	Num	4	12	12.	.	MAXIMUM KNOW AIS IN ACCIDENT
19	AISS	Num	4	12	12.	.	BODY REGION AND AIS SEVERITY
20	AINJSER	Num	4	12	12.	.	NUMBER SEVERELY INJURED OCCUPANTS
21	AINJURED	Num	4	12	12.	.	TOTAL NUMBER OF INJURED OCCUPANTS
22	ALCINV	Num	4	12	12.	.	ALCOHOL INVOLVED IN ACCIDENT
23	DRGINV	Num	4	12	12.	.	DRUG INVOLVED
24	MANCOLL	Num	4	12	12.	.	MANNER OF COLLISION
25	CRFATAL	Num	4	12	12.	.	CRASH DERIVED FATALITY

Figure 3.4: A variable view of accident table in SAS system viewer (NHTSA, 2002)

The coding of the variables can be found from the "NASS-NASS/CDS 2002-2003 Expanded Data Set Coding and Editing Manual" prepared by NHTSA (available from NHTSA's website). The codes for some of the variables used in this study are provided in Appendix A for reference.

3.1.1 Sample Selection

The sample of NASS-NASS-CDS accidents/vehicles involved for the years 2002 and 2003 are 4589/8154 and 4754/8461 respectively (refer to accident tables). The EDR pre-crash table (EDRPRECR) information includes vehicle speed (mph), engine speed

(RPM), brake-switch status and percent throttle five seconds prior to the crash at one second intervals. As EDR reports are not available for all vehicles that are involved in crashes, the EDR pre-crash information is available for only some vehicles. The actual number of records in the EDR pre-crash table are 3085 and 3970 respectively for 2002 and 2003.

SAS System Viewer - [edrprecr]

File Edit View Window Help

	SCASEID	VEHNO	RATWGT	NATWGT	EACCSEQ	STRATIF	VERSION	RECTYPE	PRESEC	SPEED	RPM	THROT	BRKSWTCH
1	12500463	2	1609.743	1.052	1	G	15	EP	-5	998	99998	998	8
2	12500463	2	1609.743	1.052	1	G	15	EP	-4	998	99998	998	8
3	12500463	2	1609.743	1.052	1	G	15	EP	-3	998	99998	998	8
4	12500463	2	1609.743	1.052	1	G	15	EP	-2	998	99998	998	8
5	12500463	2	1609.743	1.052	1	G	15	EP	-1	998	99998	998	8
6	17500449	1	677.699	0.723	1	G	15	EP	-5	42	1152	0	2
7	17500449	1	677.699	0.723	1	G	15	EP	-4	41	1152	0	2
8	17500449	1	677.699	0.723	1	G	15	EP	-3	40	1088	0	2
9	17500449	1	677.699	0.723	1	G	15	EP	-2	39	1088	0	2
10	17500449	1	677.699	0.723	1	G	15	EP	-1	7	256	0	2
11	91000000	1	1	0	1	C	15	EP	-5	998	99998	998	8
12	91000000	1	1	0	1	C	15	EP	-4	998	99998	998	8
13	91000000	1	1	0	1	C	15	EP	-3	998	99998	998	8
14	91000000	1	1	0	1	C	15	EP	-2	998	99998	998	8
15	91000000	1	1	0	1	C	15	EP	-1	998	99998	998	8
16	13400496	1	41.873	0.047	99	A	15	EP	-5	42	1536	2	2
17	13400496	1	41.873	0.047	99	A	15	EP	-4	41	1664	9	2
18	13400496	1	41.873	0.047	99	A	15	EP	-3	40	1472	2	2
19	13400496	1	41.873	0.047	99	A	15	EP	-2	40	1472	2	2
20	13400496	1	41.873	0.047	99	A	15	EP	-1	39	1280	2	1
21	12900465	1	32.452	0.026	99	A	15	EP	-5	998	99998	998	8
22	12900465	1	32.452	0.026	99	A	15	EP	-4	998	99998	998	8
23	12900465	1	32.452	0.026	99	A	15	EP	-3	998	99998	998	8
24	12900465	1	32.452	0.026	99	A	15	EP	-2	998	99998	998	8
25	12900465	1	32.452	0.026	99	A	15	EP	-1	998	99998	998	8
26	12900465	1	32.452	0.026	1	A	15	EP	-5	998	99998	998	8
27	12900465	1	32.452	0.026	1	A	15	EP	-4	998	99998	998	8
28	12900465	1	32.452	0.026	1	A	15	EP	-3	998	99998	998	8
29	12900465	1	32.452	0.026	1	A	15	EP	-2	998	99998	998	8
30	12900465	1	32.452	0.026	1	A	15	EP	-1	998	99998	998	8
31	13900496	2	288.074	0.319	1	E	15	EP	-5	15	1792	76	2
32	13900496	2	288.074	0.319	1	E	15	EP	-4	12	1408	0	2
33	13900496	2	288.074	0.319	1	E	15	EP	-3	11	1536	0	2

Ready SAS System Viewer - [edrprecr] Hdn cols:0 Obs 1 3085 of 3085

Figure 3.5: A snapshot of table showing EDR pre-crash information (NHTSA, 2002)

The number of vehicles for which the EDR data is available, i.e., the number of EDR data sets, can be determined by dividing the number of records in the EDR Precrash data table by 5, since there are 5 records for each set of pre-crash data (see above figure). The 5 records are associated with each of the 5 seconds of pre-crash data that is collected by the EDR. There are 3085/5=617 and 3970/5=794 sets of EDR pre-crash for 2002 and 2003 respectively.

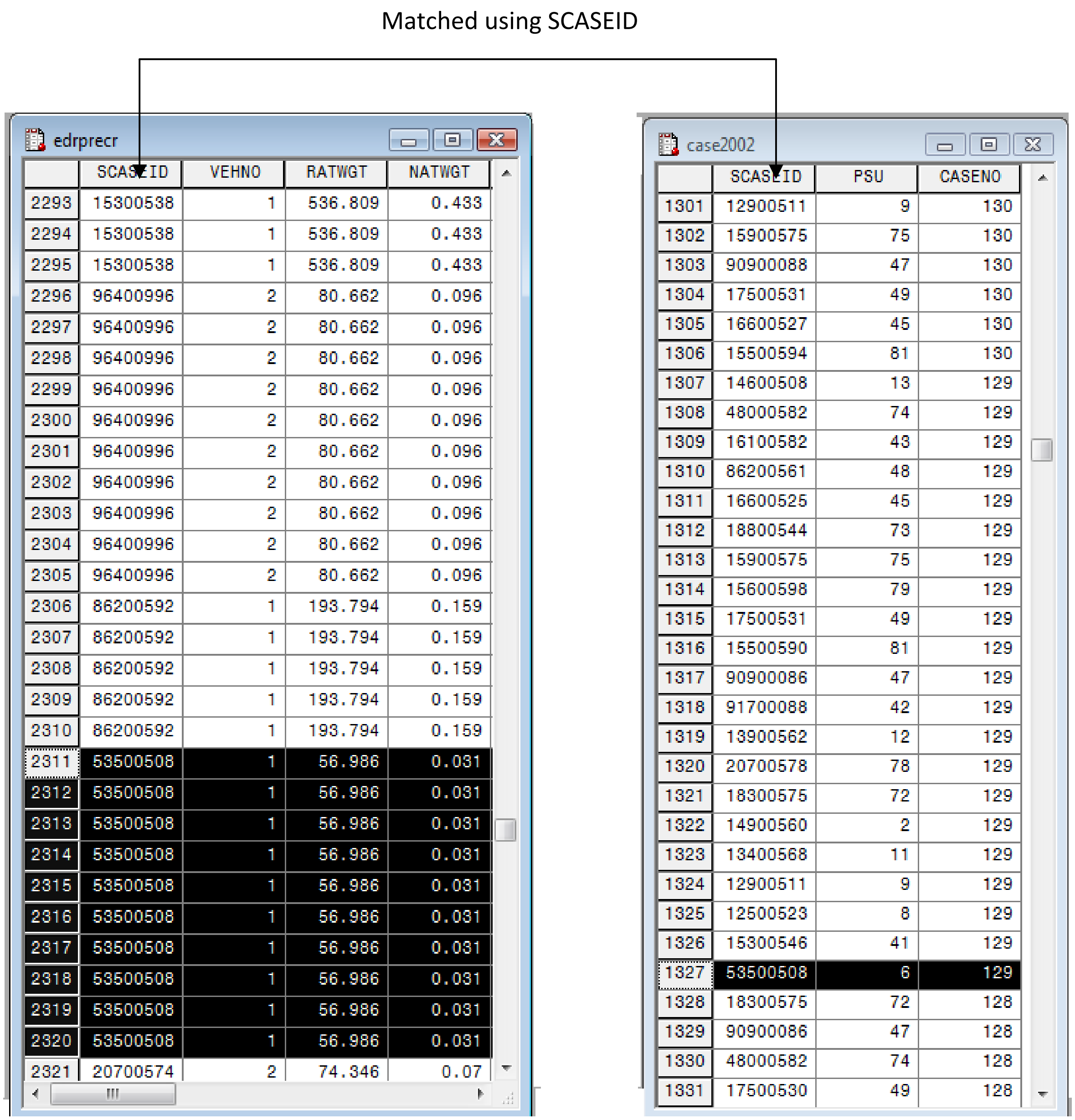
3.1.2 (a) Matching EDR Reports

The actual EDR reports that were downloaded from the crashed vehicles are available for reference or to obtain additional details about the crash. NHTSA created 2 separate SAS datasets, CASE2002 & CASE2003, which contains the mapping between the EDR Report files and the expanded NASS/CDS SAS datasets. Each dataset has the case identifier (SCASEID) and the associated sampling unit (PSU) and case number (CASENO) for that specific case. The EDR reports are saved with file names of the form:

'YEAR-PSU-CASENO-vX' (e.g. 2002-06-129-v1)

where,

- YEAR is 2002 or 2003
- PSU (Primary Sampling Unit) is an assigned city or a county for the purpose of sampling, and the field PSU reports the sampling unit that investigated the case assigned by Automated Case Selection System (ACSS)
- CASENO is the case number assigned by Automated Case Selection System (ACSS)
- vX is vehicle number (X=1, 2, 3…..)



	SCASEID	VEHNO	RATWGT	NATWGT
2293	15300538	1	536.809	0.433
2294	15300538	1	536.809	0.433
2295	15300538	1	536.809	0.433
2296	96400996	2	80.662	0.096
2297	96400996	2	80.662	0.096
2298	96400996	2	80.662	0.096
2299	96400996	2	80.662	0.096
2300	96400996	2	80.662	0.096
2301	96400996	2	80.662	0.096
2302	96400996	2	80.662	0.096
2303	96400996	2	80.662	0.096
2304	96400996	2	80.662	0.096
2305	96400996	2	80.662	0.096
2306	86200592	1	193.794	0.159
2307	86200592	1	193.794	0.159
2308	86200592	1	193.794	0.159
2309	86200592	1	193.794	0.159
2310	86200592	1	193.794	0.159
2311	53500508	1	56.986	0.031
2312	53500508	1	56.986	0.031
2313	53500508	1	56.986	0.031
2314	53500508	1	56.986	0.031
2315	53500508	1	56.986	0.031
2316	53500508	1	56.986	0.031
2317	53500508	1	56.986	0.031
2318	53500508	1	56.986	0.031
2319	53500508	1	56.986	0.031
2320	53500508	1	56.986	0.031
2321	20700574	2	74.346	0.07

	SCASEID	PSU	CASENO
1301	12900511	9	130
1302	15900575	75	130
1303	90900088	47	130
1304	17500531	49	130
1305	16600527	45	130
1306	15500594	81	130
1307	14600508	13	129
1308	48000582	74	129
1309	16100582	43	129
1310	86200561	48	129
1311	16600525	45	129
1312	18800544	73	129
1313	15900575	75	129
1314	15600598	79	129
1315	17500531	49	129
1316	15500590	81	129
1317	90900086	47	129
1318	91700088	42	129
1319	13900562	12	129
1320	20700578	78	129
1321	18300575	72	129
1322	14900560	2	129
1323	13400568	11	129
1324	12900511	9	129
1325	12500523	8	129
1326	15300546	41	129
1327	53500508	6	129
1328	18300575	72	128
1329	90900086	47	128
1330	48000582	74	128
1331	17500530	49	128

Figure 3.6: Matching the SCASEID and VEHNO to find PSU and CASENO (NHTSA, 2002)

CDR File Information

Vehicle Identification Number	1G2NW52E5XM******
Investigator	
Case Number	
Investigation Date	
Crash Date	
Filename	2002-06-129-V1.CDR
Saved on	xxxxx
Collected with CDR version	Crash Data Retrieval Tool 1.385
Reported with CDR version	Crash Data Retrieval Tool 2.900
Event(s) recovered	Deployment Non-Deployment

Figure 3.7: Matched EDR report (NHTSA, 2002)

3.1.2 (b) Data Filtering

It is to be noted that there are multiple pre-crash data sets for some vehicles. So they are to be filtered to avoid duplicates which lead to erroneous results. The coding and descriptions of the variables used for filtering are provided in the table below.

Table 3.1: Codes and descriptions of the fields used for filtering (NHTSA, n.d.)

Field	Code	Description
EACCSEQ (Event Accident Sequence)	97 98 99	Non NASS event Other, specify Unknown
INVCYCLE & EVCYCLES (Investigated Ignition Cycle & Event Ignition Cycles)	999998 999999	Not reported Unknown
SPEED (Pre-crash speed)	998 999	Not reported Unknown

A separate field ‘flag’ is created in EDR Pre-crash information table (EDRPRECR). The values for the flag are assigned based on the conditions shown in Figure 3.8. All vehicles with only one set of pre-crash speeds are coded as 0. ‘Zero’ flags are considered to be good data and are used for analysis and all non-zero flags are filtered out.

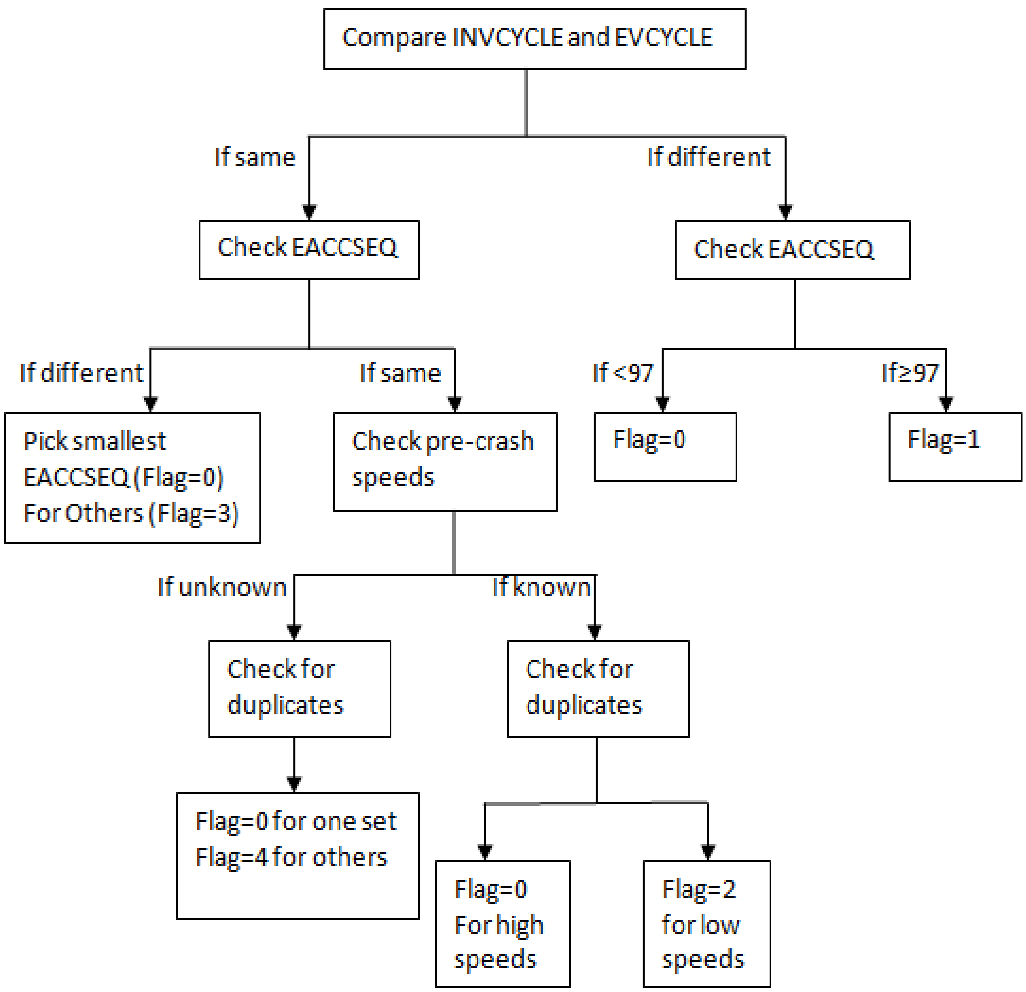


Figure 3.8: Flow chart showing conditions used for coding 'Flag'

As the interest of this study is on pre-crash speeds, the researcher assumed that the drivers chosen speed prior to a crash would be most similar to the earliest speed recorded by the EDR since drivers tend to react to looming crashes by braking. Therefore, the smallest event accident sequence number (EACCSEQ) is chosen, since it is either the first event that took place if in the case of multiple events, or the only event that took place.

The figure below shows the EDR pre-crash table with the flag field added to it. The process described above is used to populate the values of the flag field.

edr precr

flag	SCASEID	VEHNO	RATWGT	NATWGT	EACCSEQ	STRATIF	VERSION	RECTYPE	PRESEC
0	125004635	2	1609.743	1.052	1	G	15	EP	-5
0	125004635	2	1609.743	1.052	1	G	15	EP	-4
0	125004635	2	1609.743	1.052	1	G	15	EP	-3
0	125004635	2	1609.743	1.052	1	G	15	EP	-2
0	125004635	2	1609.743	1.052	1	G	15	EP	-1
0	175004491	1	677.699	0.723	1	G	15	EP	-5
0	175004491	1	677.699	0.723	1	G	15	EP	-4
0	175004491	1	677.699	0.723	1	G	15	EP	-3
0	175004491	1	677.699	0.723	1	G	15	EP	-2
0	175004491	1	677.699	0.723	1	G	15	EP	-1
0	910000001	1	1	0	1	C	15	EP	-5
0	910000001	1	1	0	1	C	15	EP	-4
0	910000001	1	1	0	1	C	15	EP	-3
0	910000001	1	1	0	1	C	15	EP	-2
0	910000001	1	1	0	1	C	15	EP	-1
0	134004967	1	41.873	0.047	99	A	15	EP	-5
0	134004967	1	41.873	0.047	99	A	15	EP	-4
0	134004967	1	41.873	0.047	99	A	15	EP	-3
0	134004967	1	41.873	0.047	99	A	15	EP	-2
0	134004967	1	41.873	0.047	99	A	15	EP	-1
1	129004654	1	32.452	0.026	99	A	15	EP	-5
1	129004654	1	32.452	0.026	99	A	15	EP	-4

Figure 3.9: The EDR precrash table with added flag field

To determine the useful cases, a query 'USEFUL CASES' is run, setting the flag value=0. By doing this, the duplicate records are removed. A crosstab named 'USEFUL CASES_Crosstab' is run for the above query, which gives the speed at -5, -4, -3, -2, -1 seconds for each vehicle in each accident. After running the filter, EDR pre-crash speed data was found to be recorded for 467 and 647 vehicles in 2002 and 2003 respectively.

Unfortunately, many of the early EDR devices did not record pre-crash speed information. Those vehicle records for which the pre-crash speeds are not known are removed. This is done by writing a subsequent query with criteria as 'not 998' for the field TOTAL OF SPEED (shown in Figure 20 generated from crosstab). The number of vehicles with EDR data and known pre-crash speeds is 249 and 385 respectively for 2002 and 2003. The total for both years includes 634 vehicles.

USEFUL CASES_Crosstab

SCASEID	VEHNO	EACCSEQ	Total Of SPE	-5	-4	-3	-2	-1
909000601	1	99	20	35	29	20	12	4
909000623	1	1	55.4	55	55	56	57	54
909000641	1	99	41	63	56	49	28	9
909000642	2	99	998	998	998	998	998	998
909000662	1	1	33.8	35	34	34	34	32
909000703	1	1	76	76	76	76	76	76
909000721	2	1	998	998	998	998	998	998
909000821	1	1	30.2	47	36	32	32	4
909000841	1	99	998	998	998	998	998	998
909000903	1	1	53.4	75	67	61	46	18
909000921	1	1	70.6	72	72	67	73	69
909000981	1	1	998	998	998	998	998	998
910000001	1	1	998	998	998	998	998	998
910000201	1	1	998	998	998	998	998	998
910000241	1	1	53.4	62	59	59	50	37
910000522	1	1	44	45	45	45	45	40
910000561	1	1	14.6	18	17	14	14	10
910000622	1	1	1.8	1	1	1	1	5
910000961	4	1	30.6	32	32	32	30	27
914000162	2	1	998	998	998	998	998	998
914000321	2	1	36.8	35	37	37	39	36
914000484	1	1	40.6	52	48	45	35	23
914000485	1	2	28.2	43	32	27	9	30
917000901	1	2	50.6	116	96	37	2	2
917000904	1	1	998	998	998	998	998	998
917000941	1	1	53.4	75	67	58	40	27
964009969	2	1	52.8	58	58	58	50	40

Record: 1 of 467 No Filter Search

Figure 3.10: Excerpt of EDRPRECR table with records involving unknown speeds highlighted in yellow

The final step in the filtration process is to determine the number of vehicles for which police reported speeds are known. This final filter is used only to determine if police reported speeds are similar to those recorded by EDRs. This filter is implemented similarly as above, except that the criterion 'not 998 or 999' is now given to travel speed (TRAVELSP). It is found that only 83 records in 2002 and 128 records in 2003 have associated police reported travel speeds. Thus, the comparison of police reported vs. EDR recorded pre-crash speeds is limited to 211 vehicles.

3.1.3 Merge Data

The analysis described in the following section requires a completed dataset with variables from several tables shown in Figure 3.11. Therefore, several queries were conducted to join USEFUL CASES_Crosstab with the general vehicle table (gv). The join property "Include ALL records from 'USEFUL CASES_Crosstab' and only those records from 'gv', where the joined fields are equal" is selected along with criteria for TOTAL OF SPEED as 'not 998 or 999' (i.e. those records with known pre crash speeds are only included). This helps to pullout the posted speed limit and also some of the other variables from the gv table (like DRINKING, LANES, ALIGNMENT, etc) that are needed for the analysis. The above mentioned process of querying (i.e., joining the USEFUL CASES_Crosstab with gv table) is repeated with other tables like driver, event etc and the other required variables (like AGE, SEX, OBJCONT etc) are selected.

The above process is explained briefly in the figure shown below.

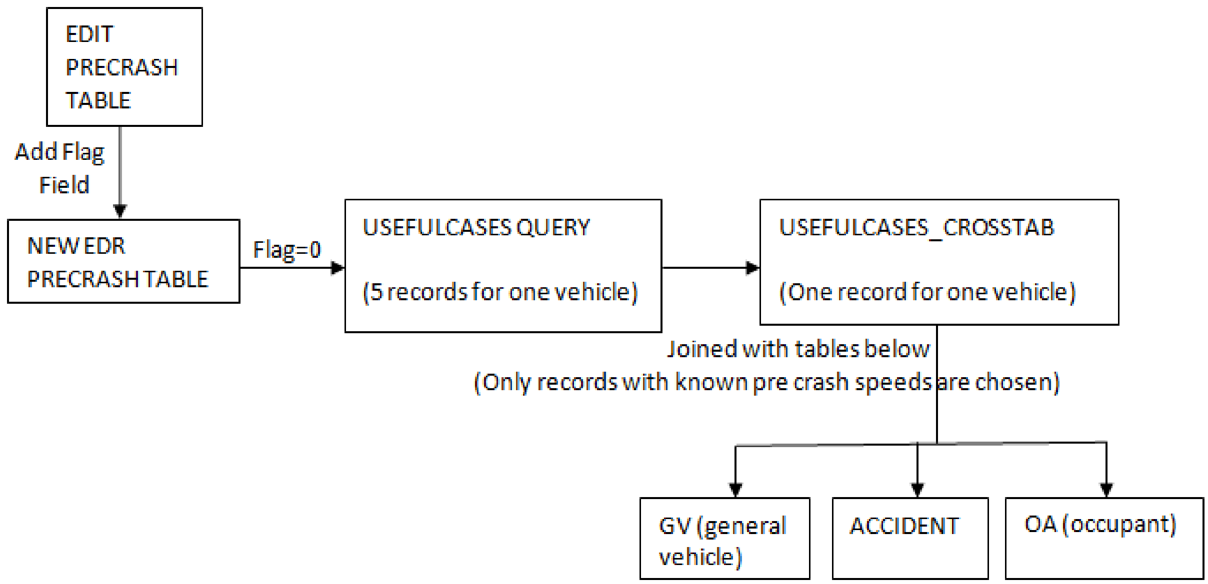


Figure 3.11: Process of extraction of data

The description and coding of various variables used in the process of querying are provided in the appendix. For example, for a query written to know the number of single vehicle crashes, the criterion for ACCTYPE is given as 'Between 1 and 10' based on coding (see appendix for coding). Also, after pulling out the required data using a query or set of queries, the final query is exported to excel and the data for the two years is combined for further analysis. It is to be noted that posted speed limit is in kmph and is to be converted to mph (1kmph=0.625 mph), as the pre-crash speeds are in mph and moreover, mph is a standard unit of speed in U.S.

3.1.4 Analysis Methods

The thorough description of the data used in this study is provided as the first step in the analysis using descriptive statistics to compare the sample to the population followed by speed trend analysis and statistical comparisons. There are three distinct sections, and each section has a specific dataset associated with it. The three sections are as follows:

- Sample Descriptive Statistics
- Speed Trend Analysis
- Statistical Comparison of EDR and Police Reported Speeds

Descriptive statistics and Trend analysis are used to understand the make-up of the subset (EDR data) with regard to the population (all crashes in the NASS/NASS/CDS database). Descriptive statistics for this analysis provide information about the frequency of crashes with respect to age and gender of driver, single vehicle and multiple vehicles, at different posted speed limit ranges. Descriptive statistics shows the comparison between the NASS/NASS/CDS and EDR data, which helps to explain if the EDR data are representative of the parent dataset (NASS/NASS/CDS). Chi-distribution tests are performed for each of these variables, with actual value as EDR data and expected value as NASS/CDS data. If the probability of independence is high, it means that EDR data is representative of NASS/CDS data, and vice versa.

In this study, the trend analysis was conducted using EDR data from all useful cases (as defined in the previous section). A subset of the NASS/NASS/CDS dataset representing only makes and models of vehicles with EDR capability was selected as the comparison group. Early analysis showed that there were some significant differences in comparisons of the EDR sample with the overall NASS/NASS/CDS population. Therefore, the researcher chose to normalize the analysis by taking only vehicles from the NASS/NASS/CDS dataset that have EDR capability. Table 5 shows the number of NASS/CDS and EDR crashes with respect to make/models. The make and model names are obtained from NHTSA's NASS/CDS coding and editing manual of 1995 (NHTSA, 1995).

Table 3.2: Make/model numbers in NASS/CDS and EDR data

MAKE	MODEL	# in EDR (634)	# in NASS (3692)	% in EDR	% in NASS
3- AM General	421- Hummer	1	2	0.2	0
18- Buick	2- LeSabre/Centurion/Wildcat,3-Electra/Electra225/Park Avenue,7-Century,10- Regal,20- Regal(FLD),401-	47	380	7.4	4.1
19-Cadillac	3- Deville/Fleetwood,5- Eldorado,14-Seville,18-	19	191	3.0	2.0
20-Chevrolet	2-Impala/Caprice,4- Corvette,9-Camaro,16-Cavalier,36-Monte Carlo,37,401-S10 Blazar 403,421- Fullsize Blazar,,431-Suburban,441-Astro Van 442-Lumina APV,471-S10/T10,481-C,K,R,V Series Pickup,482 498- Other light truck	355	1868	56.0	20.0
21-Oldsmobile	20-Cutlass,23,24,401-Bravada,441-Silhouette	28	135	4.4	1.4
22- Pontiac	2- Bonneville/Catalina/Parasienne,9-Firebird/Trans AM,16-J2000/Surbird,18-Grand AM,20-Grand Prix,401 441-Trans Sport	88	576	13.9	6.2
23- GMC	401-Jimmy/Typhoon,421-Fullsize Jimmy/Yukon,441-Safari,471-S15/T15/Sanoma 481-C,K,R,V Series Pickup	56	307	8.8	3.3
24- Saturn	1-SL,2- SC,5,7,401	39	230	6.2	2.5
59- Lexus	36	1	3	0.2	0.0

As EDR data is typically available for only one vehicle per crash, the NASS/NASS/CDS EDR capable data sample is therefore filtered further, so that there is only one vehicle per crash. A random selection process was used to identify the analysis vehicle for cases with multiple EDR capable vehicles. The final number of NASSNASS/CDS EDR capable vehicles used in the analysis is 3692 vehicles representing the same number of crashes. Further comparisons are conducted with this data and the data from useful EDR cases which were available for 634 vehicles (or crashes).

The speed trend analysis in this study provides information on typical factors related to speeding as found in previous research and data. The dataset that is used is the EDR useful cases, so there are typically 634 vehicles represented. However, some of the graphics only show speed-involved cases, which reduce the number to 278 vehicles. In this study, any vehicle traveling with speed above speed limit is considered to be speeding (usually, there will be ±7 mi/hr threshold). In other words, the number of speeding vehicles in this study are 278, and if the threshold 7 mi/hr is taken into account (PSL+7), the number of speeding vehicles are found to be only 168). Some of the factors that have been found to have relationships with speed include gender and age of driver, alcohol consumption, time of day, seat belt usage, type of road, etc. In this study, the influence of age, gender, alcohol was considered along with the type and severity of speeding-related crashes. For interpreting the sample (how well does the sample represents the population), the results are compared with known facts (like NHTSA Traffic Safety Facts).

Statistical tests are completed to test whether there is a significant difference in a single quantitative factor between two groups. Null (H_o) and alternate (Ha) hypothesis are set for each test and depending on the p-value, reject or fail to reject H_o. Three specific tests of significance have been defined below.

1) *Test whether there is a difference in average speeds of 5 seconds and 1 second prior to the crash*.

This test will be used to compare speeds between 5 seconds and 1 second prior to the crash. It is assumed that many drivers will try to brake in crash avoidance maneuvers. So if they are aware, speeds should be slower at one second prior to the crash than at five seconds prior to the crash. If this is true, it would imply that the speed at 5 seconds prior to the crash would be more indicative of the chosen operating speed of the driver.

H_o: The average difference between speed at 5 seconds prior to a crash and 1 second prior to a crash is zero

Ha: The average difference between speed at 5 seconds prior to a crash and 1 second prior to a crash is not zero

A t-test was performed for a sample of 619 speeds (since, out of 634, 17 has only either -5 or -1 speeds recorded) with 618 degrees of freedom and 95 percent confidence level. If the p-value is less than 0.05 (given the 95 percent confidence level), then the null hypothesis is rejected. That is, there will be a sufficient evidence to suggest that there is a difference in the average of speeds between 5 and 1 second prior to the crash using a significance level of 0.05. Similarly, if the p-value is greater than 0.05, the test fails to reject H_o, which means that there is no statistically significant difference in the average of speeds between speed at -5 and -1 seconds.

To fulfill the objective to check if police reported speeds are same as EDR obtained speeds, the same t-test as above is performed. If they are proved to be different, further tests are done to find the correlation between these speeds.

The same test for police reported speeds is repeated with the speeds at 5 seconds prior to the crash and speeds at 1 second prior to the crash. In other words, a t-test is performed to find a) if the speed at 5 seconds prior to a crash reported by EDR is same as police reported speed; and/or b) if the speed at one second prior to the crash reported by EDR is same as police reported speeds.

A sample of 210 crashes is used for these tests, because only 210 of 619 crashes contained both police reported speeds and EDR recorded speeds. As described above, the p-value is used in making a decision of rejecting and accepting H_o. The null and alternate hypotheses for these tests are as follows:

For 2(a):

H_o: The average difference between speed at 5 seconds prior to a crash and police reported speed is zero

Ha: The average difference between speed at 5 seconds prior to a crash and police reported speed is not zero

For 2(b):

H_o: The average difference between speed at 1 second prior to a crash and police reported speed is zero

Ha: The average difference between speed at 1 second prior to a crash and police reported speed is not zero

If there is a significant difference between the police reported speed and EDR recorded speed, a test is performed to determine if there is a correlation between the police reported speed and the EDR speed (either speed at-5 or speed at -1) that is representative of the chosen speed.

3) *To test whether there is a correlation exists between police reported travel speeds and EDR obtained pre-crash speeds.*

H_o: There is no correlation between police reported travel speed and EDR obtained pre-crash speeds (-5 and -1)

Ha: There is a correlation between police reported travel speed and EDR obtained pre-crash speeds (-5 and -1)

As mentioned earlier, p-value is used to decide whether to accept or to reject H_o. Pearson Correlation Coefficient (r) is used to know the magnitude and strength of correlation. The value of r ranges from -1 to +1. If r is positive, there exists a positive correlation between the tested speeds and if r is negative, there exists a negative correlation between the tested speeds. If r value is close to -1 or +1, it means that the speeds are highly correlated. If r value is close to zero, it shows that the tested speeds are not highly correlated.

3.1.5 Data Analysis

The data extracted from above is analyzed using the methods described. It should be noted that this research uses the number of crashes regardless of number of vehicles involved in the crashes. The coding used to extract the required values of the data fields are provided in the appendix. The results from the data analysis are shown in Chapter 4.

CHAPTER 4

ANALYSIS AND RESULTS

The results of this study are divided and shown in three sections. The first section provides some descriptive statistics for the sample of data. The second section provides information on how the EDR data compares to long-standing trends found in the national data, and also provides some comparisons of speeding and non-speeding crashes. Recall that, speeding crashes are found from the crash data using the posted speed limit and speed at -5 seconds. In other words, if the driver travels at a speed (speed at -5 seconds) greater than posted speed limit, that crash is considered as a speed-related crash. The third section presents the results for the statistical hypothesis tests.

4.1 Descriptive Statistics

As it is important to understand the subset of population, the crash data is analyzed to determine if the sample and population have the same characteristics. To do this, a special sample of EDR capable vehicles was located within the NASS/NASS/CDS dataset. While these vehicles are of the same make and models as those with reported EDR data, these EDR data from these vehicles were not obtained. The general statistics of crashes and the results are provided below. A Chi-distribution test is performed for each of the variables (p values are shown in brackets under each graph) and depending on the result, it is statistically confirmed whether or not the EDR data is representative of the NASS/CDS data.

Figure 4.1 shows the relative involvement of drivers based on gender. It is observed that males are involved in more crashes than females. This is the same with

NASS/NASS/CDS data but, when the NASS/NASS/CDS and EDR results are compared, it is observed that the percentage of male drivers is less and percentage of female and unknown drivers is more in the EDR data relative to NASS/NASS/CDS data.

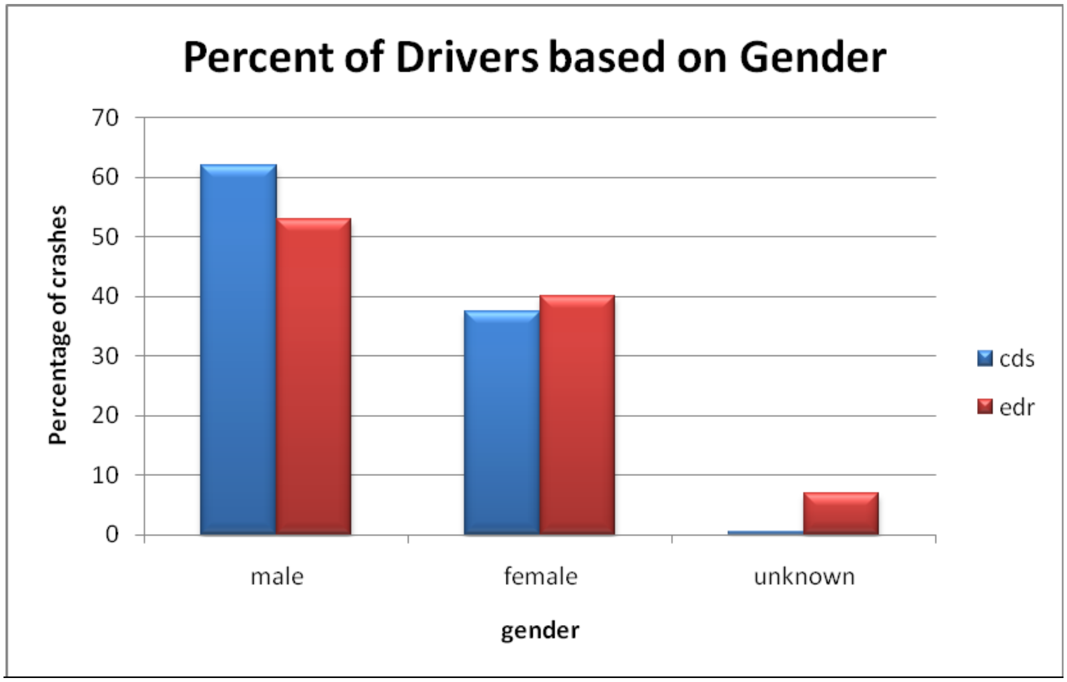


Figure 4.1: Percent of drivers based on gender with respect to NASS/NASS/CDS and EDR data (p=0)

Figure 4.2 shows the distribution of crashes with respect to drivers in different age groups. When compared to the results from the NASS/NASS/CDS data, it is observed that there is a pattern of fewer drivers in EDR dataset in younger ranges and more drivers in older ranges. Given the required consent process to obtain the EDR data, this slight bias may be a result of the higher level of difficulty of obtaining minor consent. The chi-squared distribution test (p-value 0.817) indicates that it is highly probable that the differences in the two datasets are random.

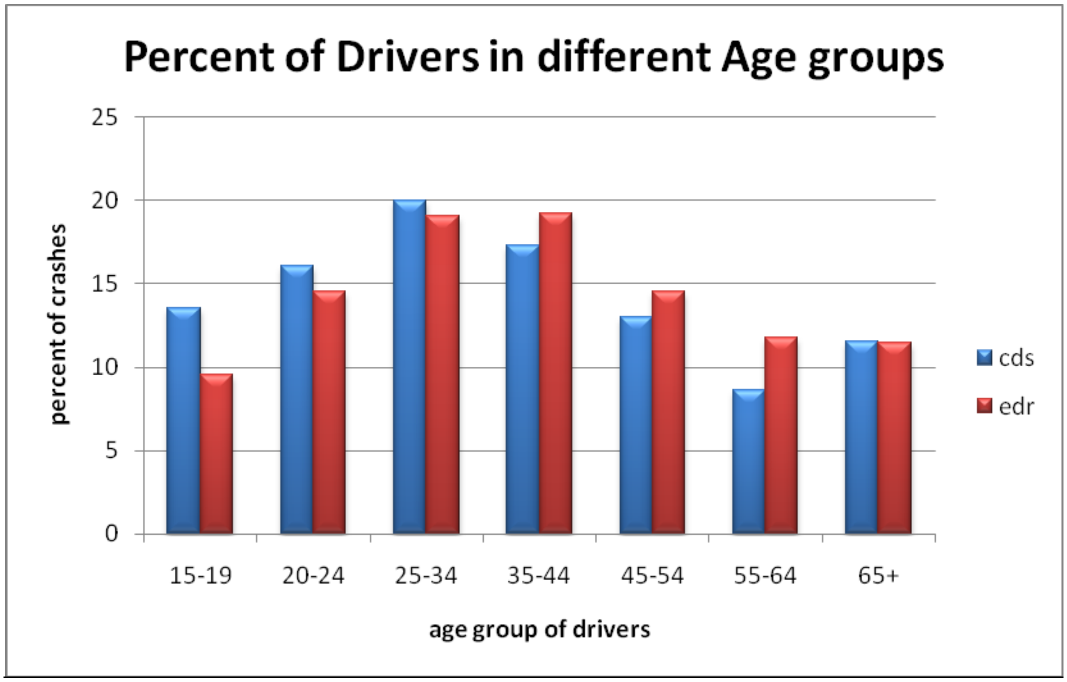


Figure 4.2: Percent of drivers in different age groups with respect to NASS/CDS and EDR data (p=0.817)

Figure 4.3 shows the distribution of crashes with respect to time of day. It is observed that, the results from both NASS/CDS and EDR data shows high frequency of crashes between 3-6 pm. But in case of time with fewer crashes, EDR shows midnight-3 am where as NASS/CDS shows 3- 6am. Figure 4.4 shows the distribution of crashes with respect to day of week. The results from the NASS/CDS and EDR data showed that the frequency of crashes is high on Friday and low on Sunday.

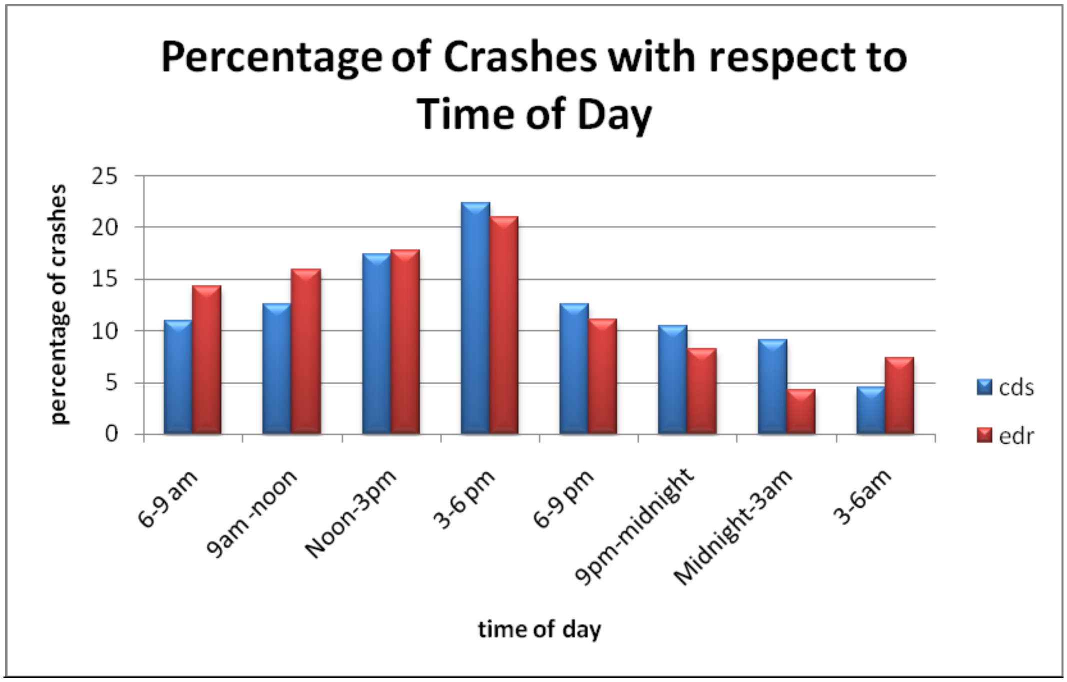


Figure 4.3: Percentage of crashes with respect to time of day (p=0.4265)

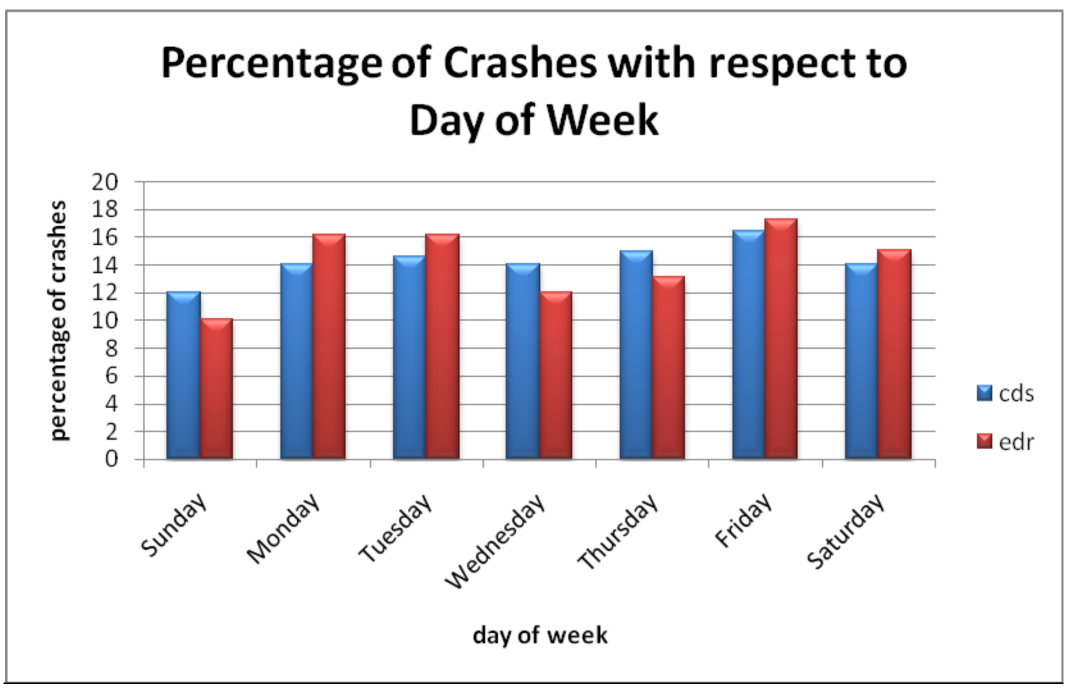


Figure 4.4: Percentage of crashes with respect to day of week (p=0.960544)

Figure 4.5 shows the percentage of speed-related and not speed-related crashes obtained from NASS/NASS/CDS and EDR data. It is interesting to note that these percentages (speeding and not speeding) are exactly opposite. In other words, the percentage of speeding and not speeding crashes from NASS/CDS data is 57 and 43, where as these percentages from EDR data are 43 and 57 respectively. The chi-squared distribution test also indicates that these two datasets are independent in this regard.

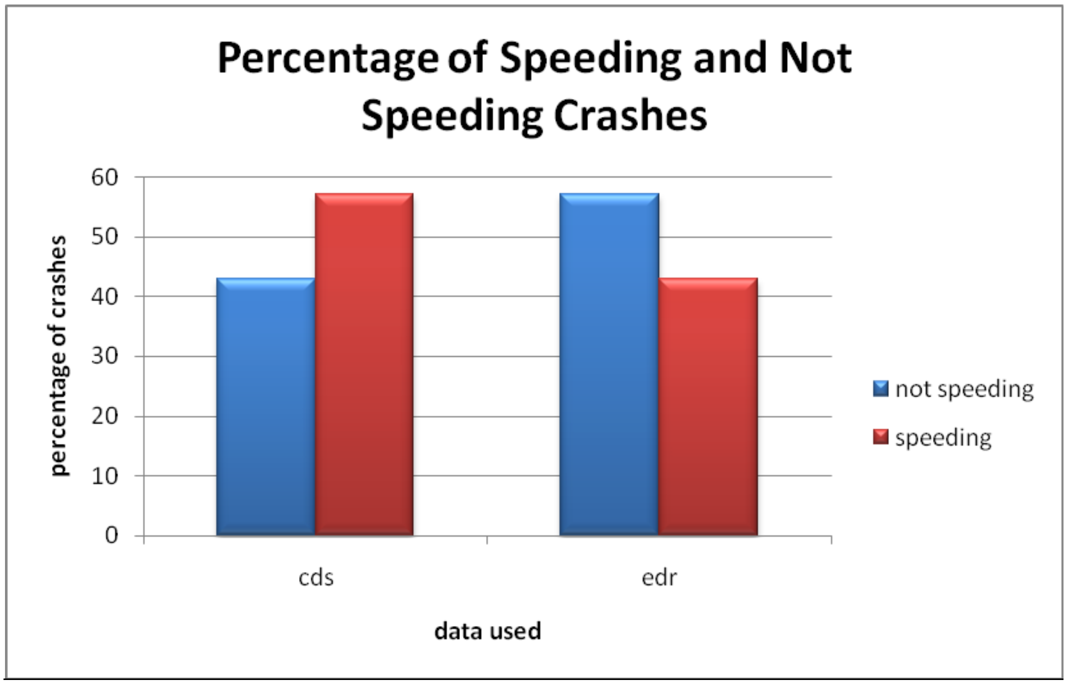


Figure 4.5: Percentage of speeding and not speeding crashes with respect to data used (p=0.0046)

The percent of drivers involved in crashes, with influence of alcohol and drugs at the time of crash are shown in Figure 4.6. NASS/NASS/CDS data showed that 10.7% of drivers are under the influence, whereas the EDR data showed that only 6.3% of the drivers are under the influence. A chi-square distribution test cannot be performed for only two variables; however, the number of drivers in each data set should provide significance. As compared to the overall NASS/NASS/CDS dataset including all makes and models, the difference was even more severe at EDR (40 out of 634 and 6.3%) and NASS/NASS/CDS (381 out of 3552 and 10.7%). Thus, it is likely that these datasets are comparable.

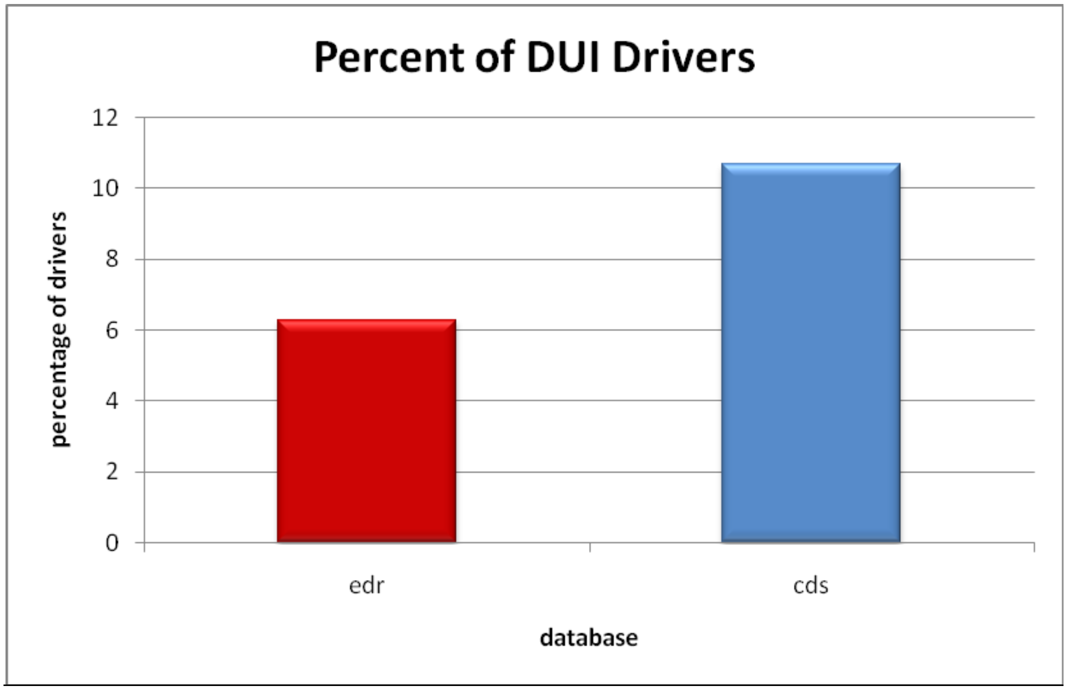


Figure 4.6: Percent of DUI drivers with respect to NASS/NASS/CDS and EDR data

Figure 4.7 shows the percentage of crashes with respect to accident type obtained from NASS/CDS and EDR data. It is interesting to observe that the results from both datasets are nearly identical. This is also indicated in the high p-value for the chi-test. Figure 4.8 shows the injury severity levels of people involved in crashes obtained from NASS/CDS and EDR data. EDR data showed a smaller number of uninjured when compared to NASS/CDS data and a higher number of injured persons. The low chi-test probability confirms the difference.

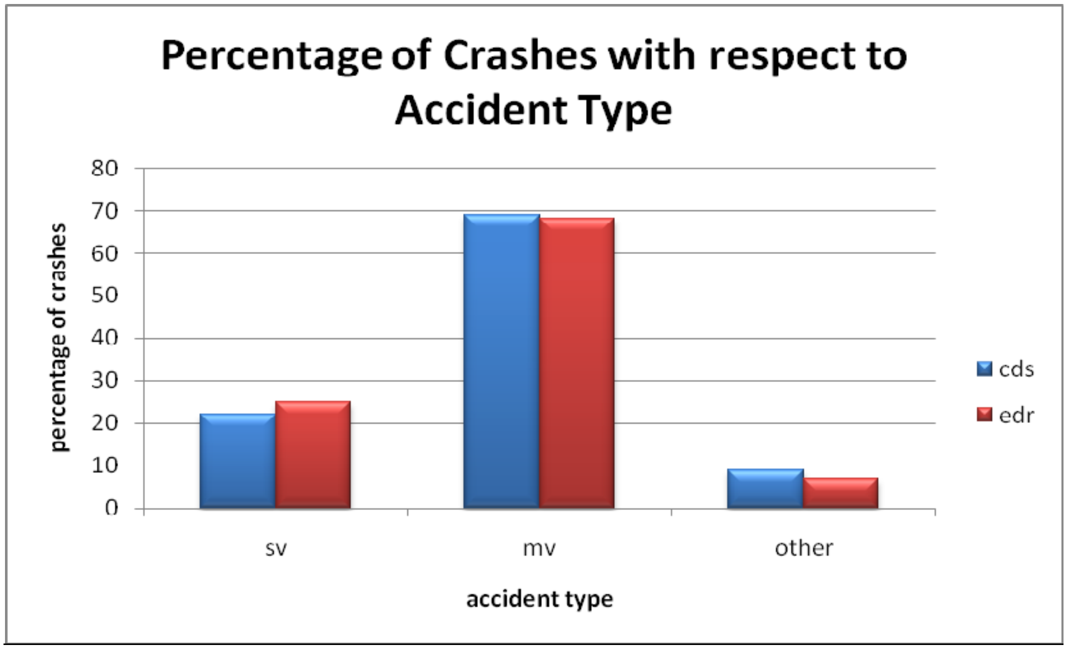


Figure 4.7: Percentage of crashes with respect to accident type (p= 0.65)

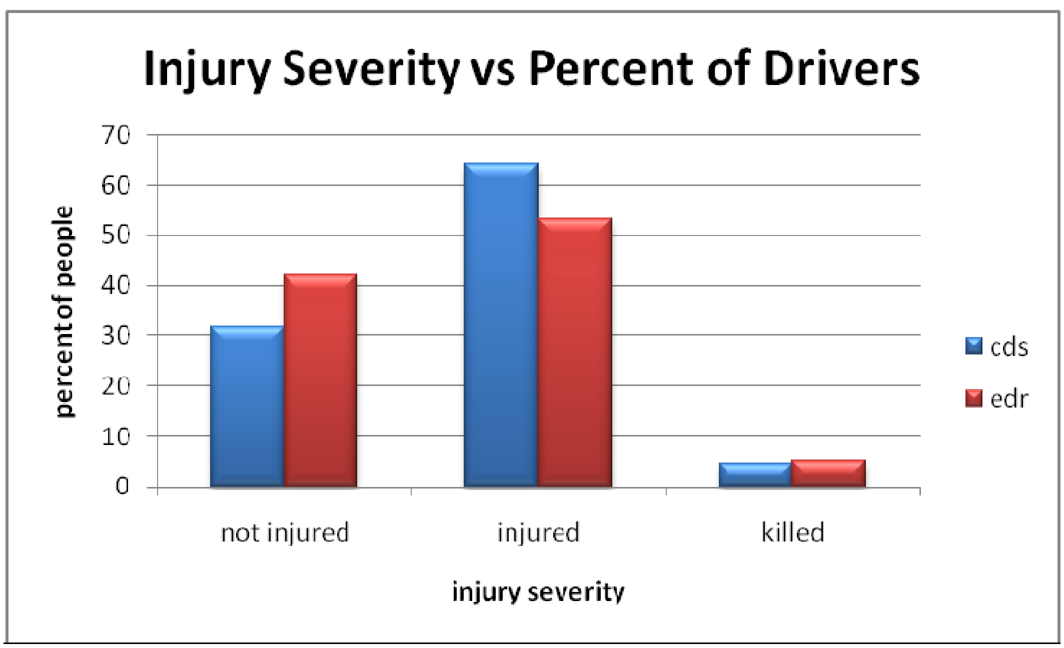


Figure 4.8: Injury severities of drivers involved in crashes with respect to NASS/CDS and EDR data (p=0.0657)

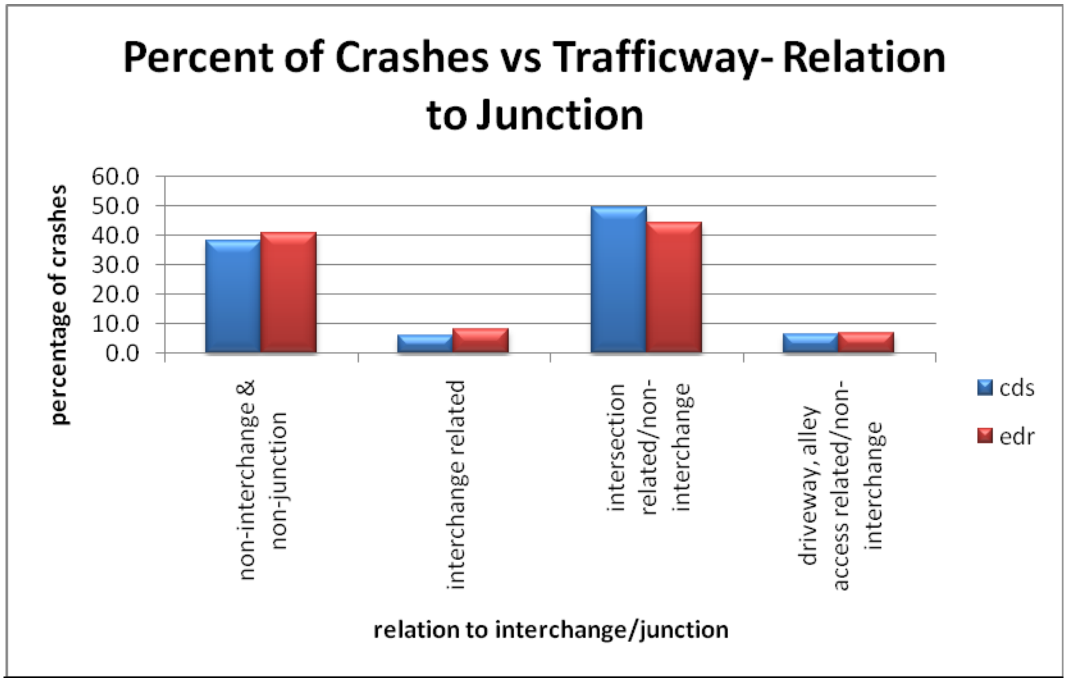


Figure 4.9: Crash frequency with respect to traffic way relation to junction (p=0.6959)

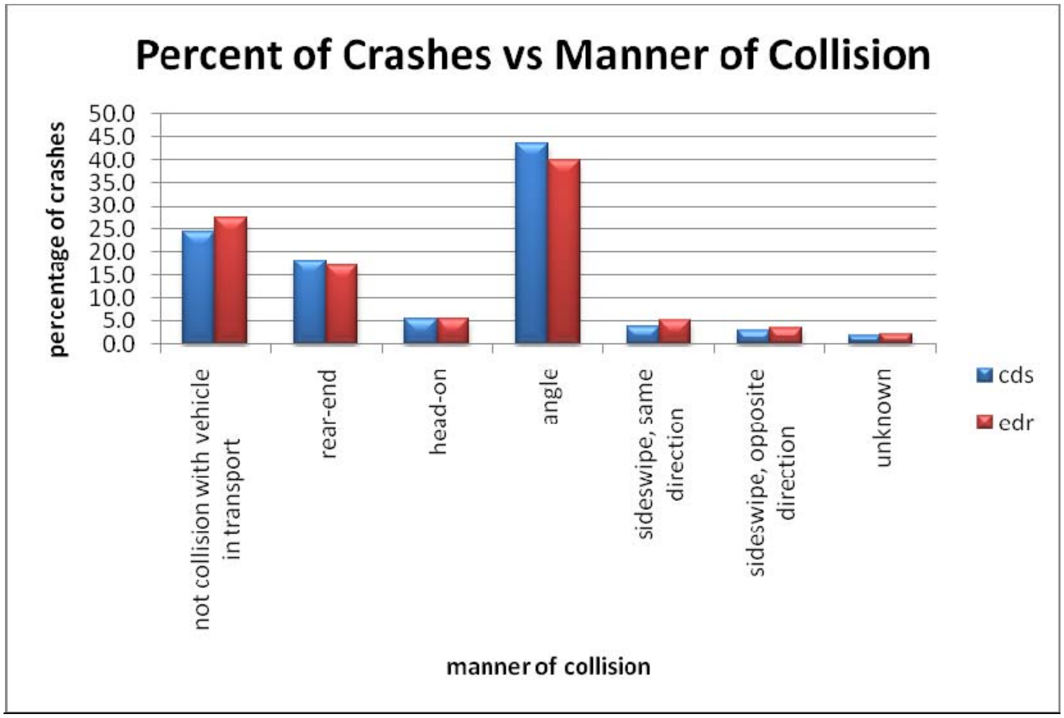


Figure 4.10: Percentage of crashes with respect to manner of collision (p=0.9654)

Figures 4.9 and 4.10 show the percent of crashes with respect to relation to junction and manner of collision respectively. From charts, it is observed that there are only slight differences in the results from NASS/CDS and EDR data, and the chi-test results showed that the EDR data is representative of NASS/CDS data.

4.2 Trend Analysis Results

Section 2 gives the findings on speeding-related crashes. This section uses on the EDR useful cases data from NASS/CDS. There are 619 crashes represented. The figure below gives the frequency of crashes in which the drivers are speeding and not speeding. It is observed that 45% (278 crashes) involve speeding which is a serious issue of concern especially if the magnitude of speeding is high. In consideration of the typical operating speeds observed by Fitzpatrick et al. (year), a 7 mi/hr threshold above the posted speed limit was also to determine its affects on the overall determination of speed related crashes. With the revised speeding threshold, the speeding crashes were still found to be 35% of all EDR crashes – see Figure 4.11. This is still a large number of crashes involving speed. For comparative purposes, some analysis results are shown by the relative comparison of speeding and non-speeding crashes.

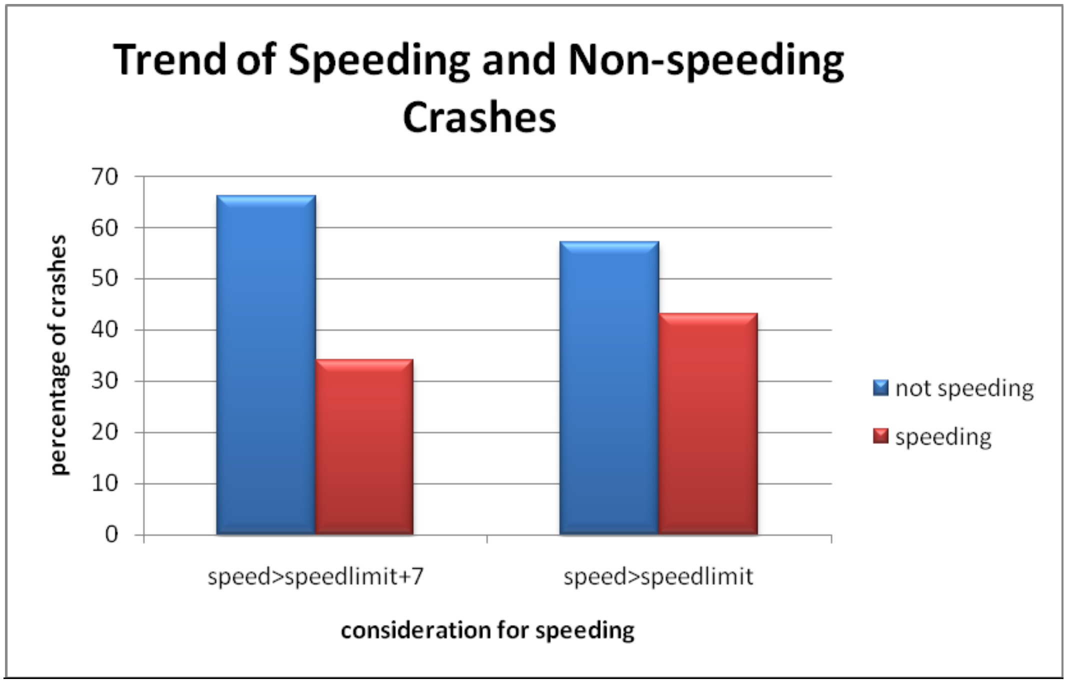


Figure 4.11: Trend of speeding and non-speeding crashes

For better understanding of speed and non-speeding crashes, this information was grouped by posted speed limits (Figure 4.12). The frequency of speeding crashes is higher on the roadways with speed limits less than 25 mi/hr (which is dangerous as most of them are residential areas). Speed-related crashes are also slightly higher on roads with high speeds to begin with – which becomes a deadly combination.

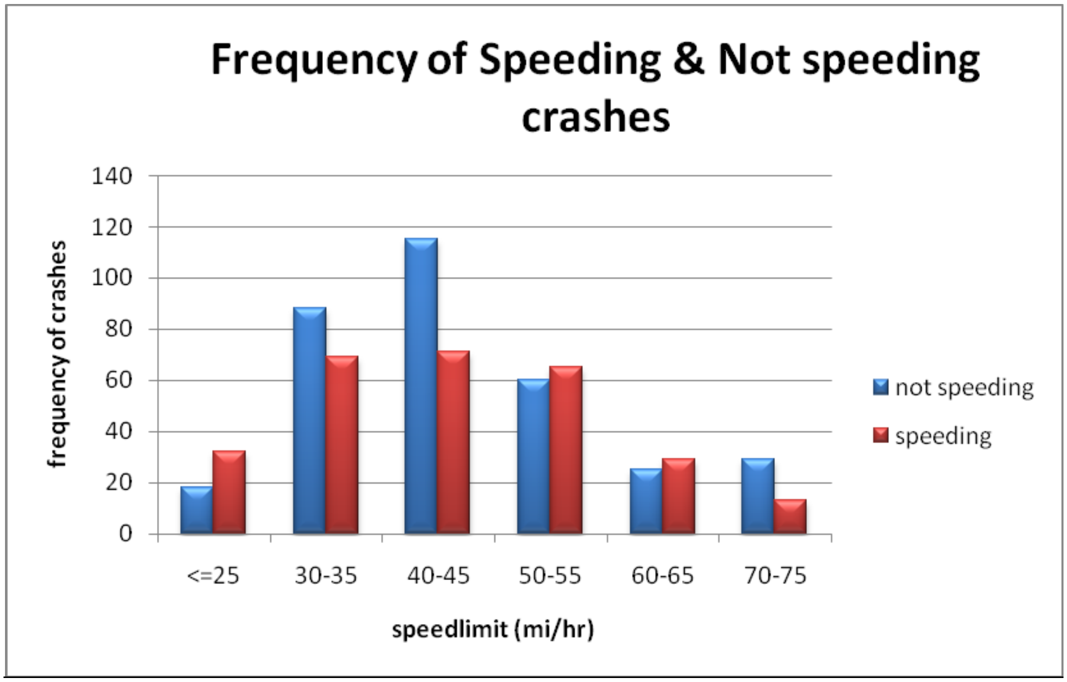


Figure 4.12: Frequency of speeding and non-speeding crashes at various speed limits

Previous research has shown that involvement in speed-related crashes varies by the age and sex of the driver. Figure 4.13 represents the frequency of speeding and non-speeding crashes based on gender. Males are clearly involved in more speed related crashes than their female counterparts.

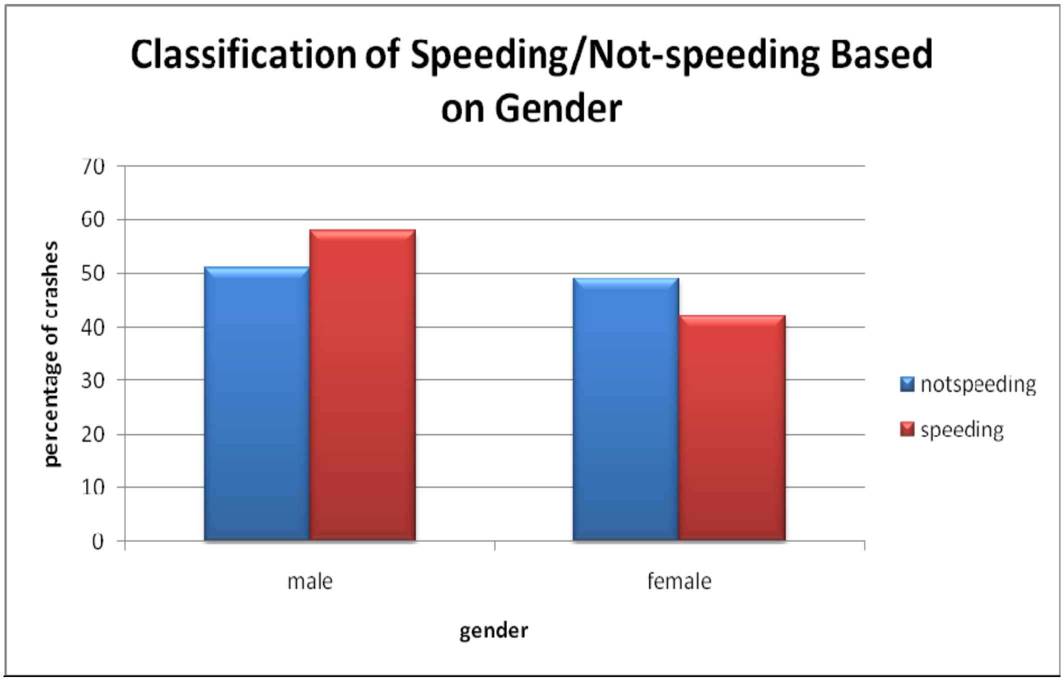


Figure 4.13: Speeding and non-speeding crashes based on gender

Figure 4.14 shows the frequency of speeding and non-speeding drivers by age group. The drivers in age group 20-24 have a higher occurrence of speeding than all the other groups. Also, it is observed that the drivers of age between 20 and 44 are speeding in more crashes than that of the other ages.

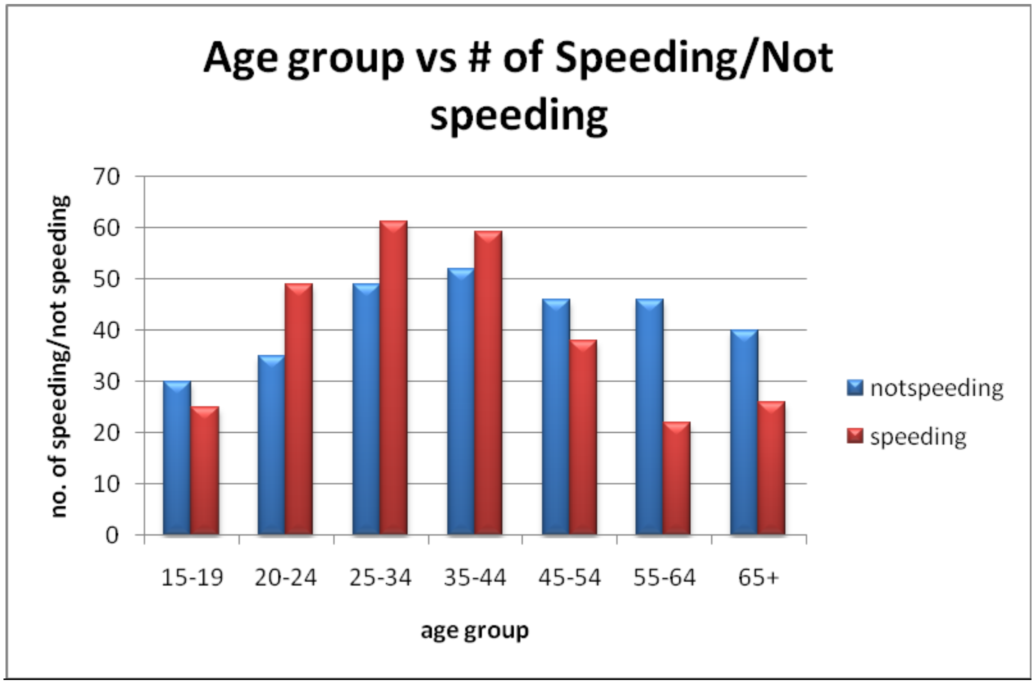


Figure 4.14: Number of speeding and non-speeding drivers in different age groups

To be clearer, it is important to know the amount of speeding by drivers in each age group. Figure 4.15 shows the amount of speeding or over speeding (how much above the speed limits are the drivers traveling) of drivers in different age groups. According to Fitzpatrick et al (2003), speeding up to 7 mi/hr is typical on most facility types. That is, almost all travel occurs at speeds above posted speed limit. But Figure 4.15 indicates that younger drivers (20-25 years of age) are more likely to exceed the speed limit by greater amounts than their older counterparts.

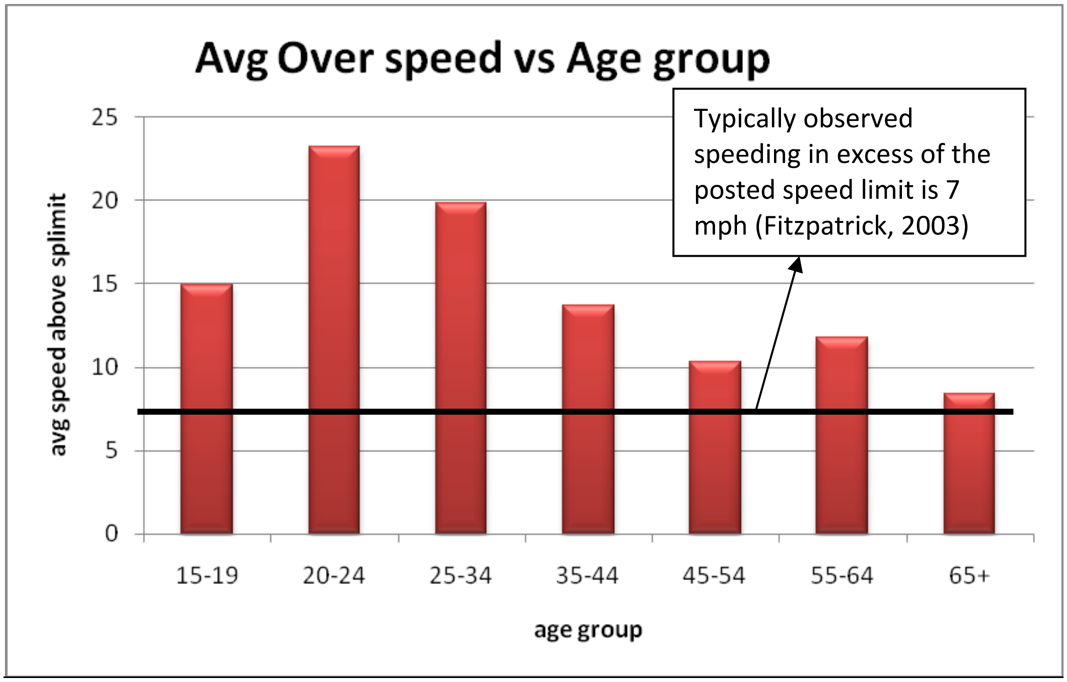


Figure 4.15: Amount of speeding by drivers in different age groups

Another popular finding from previous research included alcohol and drugs as contributing factors in speed related crashes. Figure 4.16 shows the number of crashes involving drunk or drugged drivers and whether or not they were speeding. In other words, it shows the number of drunk drivers that are found to be speeding. Out of 40 drivers that were found to be under the influence of drugs or alcohol, 30 were also speeding. While drunk driving and speeding are highly correlated, it is important to note that only 30 of 286 involved alcohol. So, speeding is an issue of general importance.

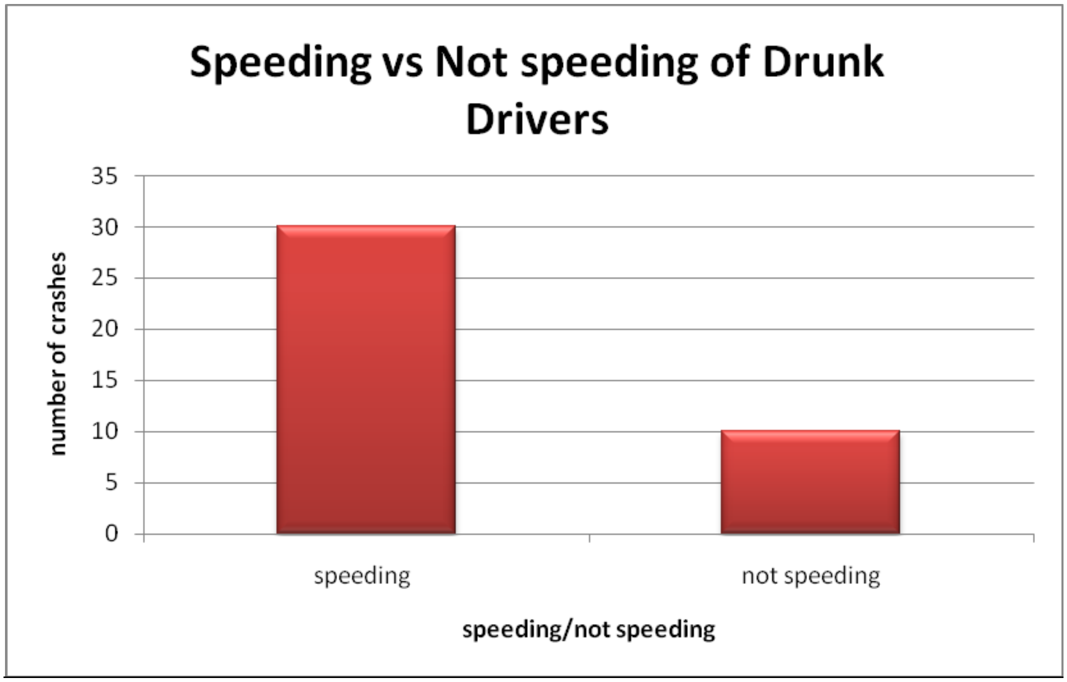


Figure 4.16: Number of drunk drivers speeding and not speeding

The average amount of speeding above the posted speed limit by drivers under the influence is shown in the figure below. It can be observed that all of the drivers under the influence are traveling at speeds much higher than the posted speed limit, and much faster than speeders who were not under the influence of drugs or alcohol.

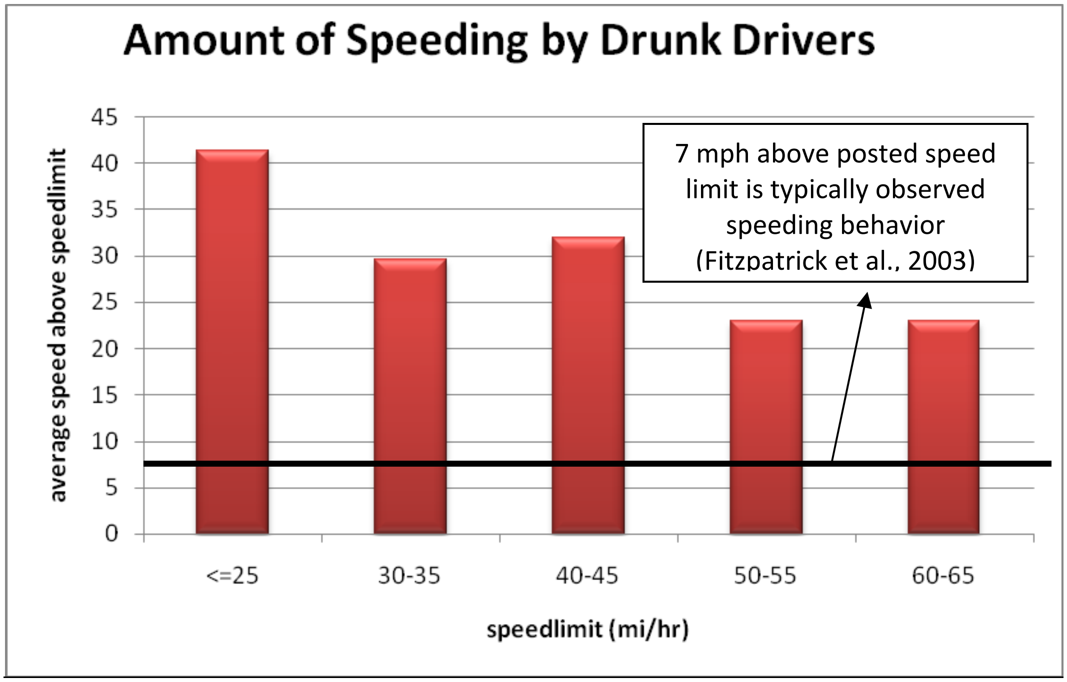


Figure 4.17: Average amount of speeding by drunk drivers

Figure 4.18 shows the frequency of speeding and non-speeding crashes with different injury severity levels. It is observed that more speeding related crashes involve moderate or severe injuries; where as non-speeding crashes are more likely to involve minor or no injuries.

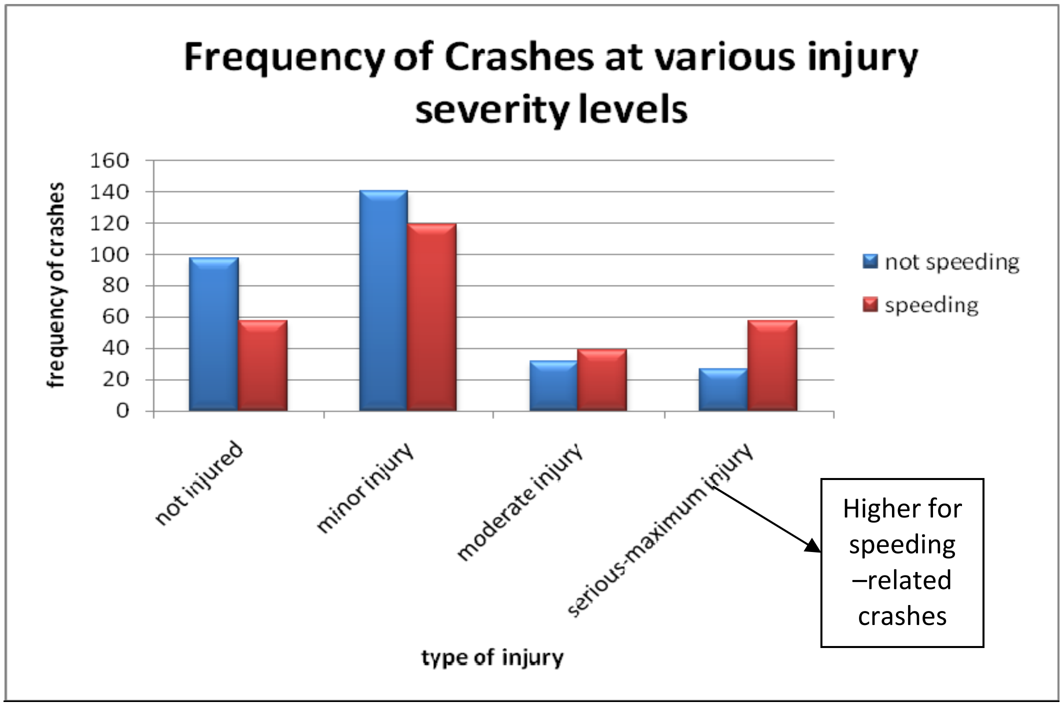


Figure 4.18: Frequency of crashes with respect to type of injury

Speeding has also been a traditional factor in single vehicle run-off road crashes on higher speed roadways. This holds true in Figure 4.19 which show the relative frequency of speed-related crashes for single vehicle and multiple vehicle crashes at various posted speed limits. Multiple vehicle crashes are much higher than single vehicle crashes at speed limits of 30-50 mph which are more common for roadways that provide high levels of access where vehicle paths are likely to cross.

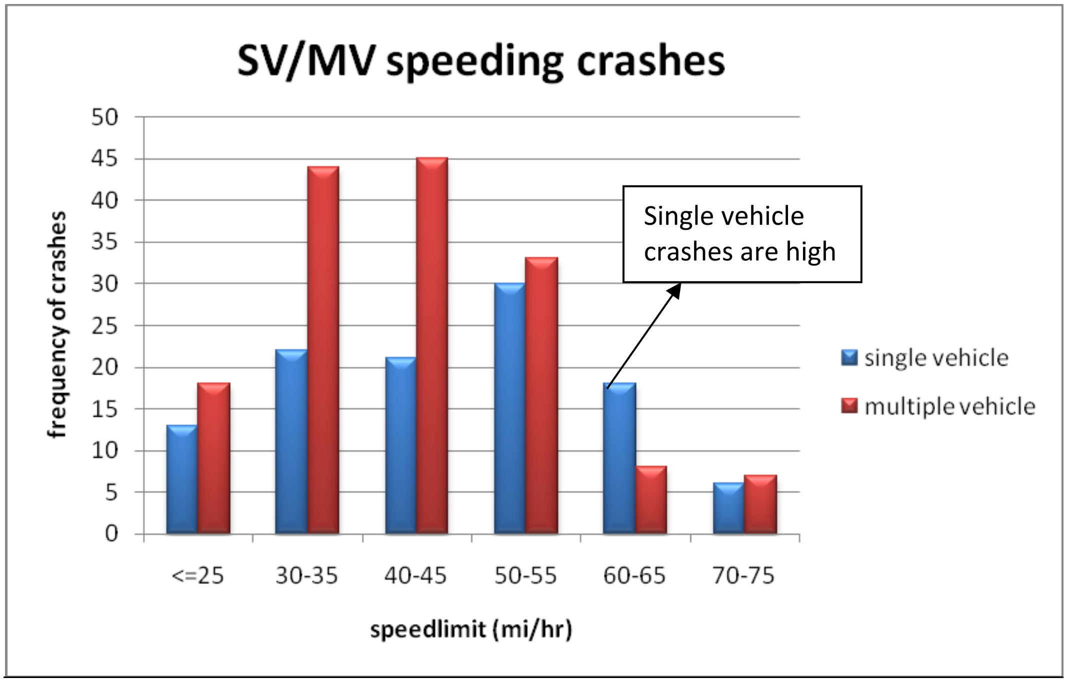


Figure 4.19: Frequency of speeding crashes of different crash types

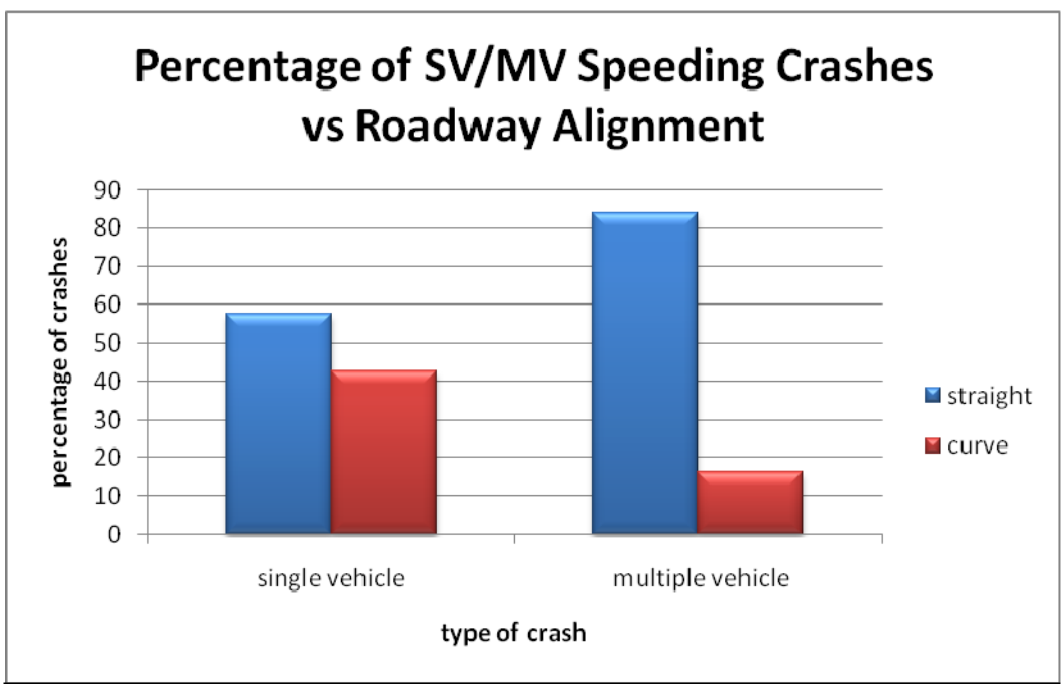


Figure 4.20: Percentage of SV/MV crashes with respect to roadway alignment

Figure 4.20 shows the percentage of single vehicle and multiple vehicle speed-related crashes on straight and curved roadway sections. It is observed that the frequency of speed-related crashes on straight section of road is high when compared to that of a curved section of road. Another important finding is that the single vehicle seeding crashes are high on curved sections when compared to that of multiple vehicle speeding crashes.

4.3 Statistical test results

Statistical tests are completed to test whether there is a significant difference in a single quantitative factor between two groups. Null (H_o) and alternate (Ha) hypothesis are set for each test and depending on the p-value, reject or fail to reject H_o. The sample size used and the results of the three specific tests of significance have been presented below.

1) *To test whether there is a difference in the average of speeds at 5 seconds and 1 second prior to the crash.*

H_o: The average difference of speeds at 5 and 1 seconds prior to the crash is zero

Ha: The average difference of speeds at 5 and 1 seconds prior to the crash is not zero.

The descriptive statistics of the data used and results obtained for this test are shown in Table 4.1 .

Table 4.1: Test for speeds at 5 and 1 seconds prior to the crash

Measure	Value
Sample Size	619
Mean of speed at -5 seconds	40.38 mi/hr
Mean of speed at -1 seconds	26.67 mi/hr
Mean Difference	13.722 mi/hr
Significance Level	5%
P-Value	<0.0001
Significant	Yes

P-value obtained for this test is <0.0001. As this is less than 0.05 (given a 95 percent confidence level), the null hypothesis is rejected. That is, at 95 percent confidence level, there is a sufficient evidence to prove that there is a significant difference in the average of speeds at five and one seconds prior to the crash.

2) *To test (a) if the speed at 5 seconds prior to a crash reported by EDR is same as police reported speed; and/or (b) if the speed at one second prior to the crash reported by EDR is same as police reported speeds*.

The null and alternate hypothesis for (a) are as follows.

H_o: The average difference between speed 5 seconds prior to a crash and police reported speed is zero

Ha: The average difference between speed 5 seconds prior to a crash and police reported speed is not zero

The statistics of the data used and the results obtained are shown in the table below.

Table 4.2: Test for police reported speed and speed at -5 seconds

Measure	Value
Sample Size	210
Mean of speed at -5 seconds	38.3 mi/hr
Mean of police reported speed	35.2 mi/hr
Mean Difference	3.1 mi/hr
Significance Level	5%
P-Value	<0.0001
Significant	Yes

Therefore, at a 5% significance level, there is sufficient evidence to suggest that there is a significant difference in the average of police reported speed and speed 5 seconds prior to the crash.

The following are the null and alternate hypothesis for (b):

H_o: The average difference between speed 1 second prior to a crash and police reported speed is zero

Ha: The average difference between speed 1 second prior to a crash and police reported speed is not zero

The table below provides the statistics of the data used and the results obtained.

Table 4.3: Test for police reported speed and speed at -1 second

Measure	Value
Sample Size	210
Mean of speed at -1 second	24.4 mi/hr
Mean of police reported speed	35.2 mi/hr
Mean Difference	-10.8 mi/hr
Significance Level	5%
P-Value	0.0174
Significant	Yes

So, it is concluded that there is a significant difference in the average of speed 1 second prior to the crash and police reported speed.

From the statistical tests, it is known that the speed 5 seconds prior to the crash is a better indicator of chosen speed and speed 1 second prior to the crash is not as reliable for pre-crash estimate. It is proved that there is a significant difference between speeds recorded by EDR and police reported speeds. At the same time, the difference in the average of speed 5 seconds prior to the crash recorded by EDR and police reported speed is 3.1 mi/hr which is practically acceptable.

The correlation between the speed 5 seconds prior to the crash and police reported speed can be seen in Figure 4.21. The p-value for the test performed to know whether there is a correlation between police reported speed and speed at -5 seconds is <0.05. So there exists a correlation between these two speeds. The Pearson Correlation Coefficient r=0.743 shows that these speeds are positively correlated and that there exists a good correlation between them.

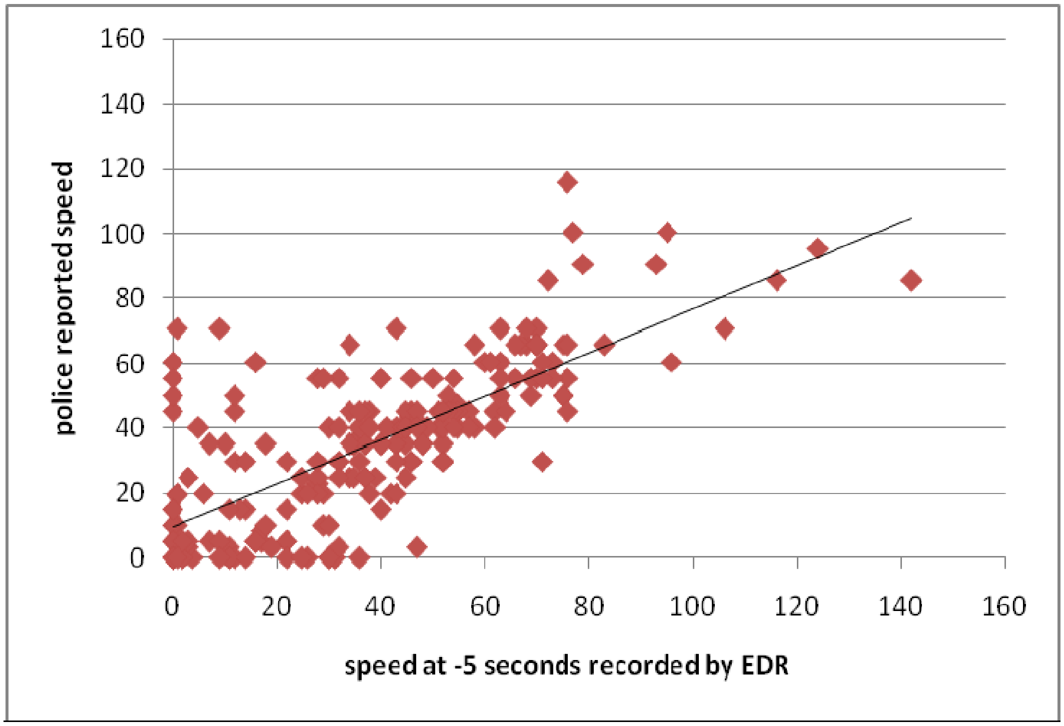


Figure 4.21: Correlation between speed at -5 seconds and police reported speed

CHAPTER 5

CONCLUSIONS AND FUTURE RECOMMENDATIONS

With ‘speed’ accounting for more than 31 percent of all traffic fatalities and costing around $40.4 billion, it is understood that speed is a serious issue of traffic safety, and that the speed and crash data obtained are prone to error. Event Data Recorder (EDR) data (pre-crash, crash and post crash information) is been collected and stored by National Highway Traffic Safety Administration (NHTSA) to overcome this issue of data accuracy and reliability, and to improve traffic safety.

The goals of this research are as follows:

- Check if EDR subset match NASS
- Determine the trend in speed crashes
- Determine if police reported speeds match speeds recorded by EDRs and/or determine if any correlation exists between the two.

For this study, a sample of 634 crashes from the years 2002 and 2003 that have complete EDR information is used. Data of corresponding variables pulled out using queries written in MS Access are used for the analysis. Descriptive statistics, trend analysis and statistical tests are done to fulfill the above mentioned goals.

The descriptive statistics of the EDR data are compared with that of the NASS/NASS/CDS data, which showed that the EDR subset has some bias that is not apparent in overall NASS sample. This bias can be clearly understood from the differences in results obtained from NASS/CDS and EDR data. For example, the percent of DUI drivers in crashes is found to be 6.3% from NASS/CDS data, where as it is 10.7% from EDR data. Also, there are differences with respect to time of day, age and sex of drivers, etc which can be clearly observed from the p values obtained from the chi tests.

The trend analysis of speeding crashes resulted in the same trend as before, which can be confirmed from the well known facts like males are more involved in speeding crashes than females, severity of crashes increases with increase in speeding, amount of speeding is high in case of DUI drivers, etc.

It was proved from the statistical test that speed 5 seconds prior to the crash is representative of the chosen speed and speed 1 second prior to the crash is not reliable for the pre-crash speed estimate. It was also proved that there is a significant difference in the average of police reported speed and speed recorded by EDR. Further, it is observed that the difference in the average of police reported speed and speed 5 seconds prior to the crash recorded by EDR is practically acceptable.

Finally, the analysis would be much effective if the EDR data is available for all the vehicles that were involved in crashes and, at the same time if the data is clear and free of duplicates.

APPENDICES

Appendix A: Coding of Variables Used in Analysis (NHTSA, 2002)

Crash Configuration: Forward Impact
34 This Vehicle's Frontal Area Impacts Another Vehicle
35 This Vehicle Is Impacted by Frontal Area of Another Vehicle
36 This Vehicle's Frontal Area Impacts Another Vehicle
37 This Vehicle Is Impacted by Frontal Area of Another Vehicle
38 This Vehicle's Frontal Area Impacts Another Vehicle
39 This Vehicle Is Impacted by Frontal Area of Another Vehicle
40 This Vehicle's Frontal Area Impacts Another Vehicle
41 This Vehicle Is Impacted by Frontal Area of Another Vehicle
42 Specifics Other
43 Specifics Unknown

Crash Configuration: Sideswipe/Angle
44 Straight Ahead on Left
45 Straight Ahead on Left/Right
46 Changing Lanes to the Right
47 Changing Lanes to the Left
48 Specifics Other
49 Specifics Unknown

Crash Category: Same Trafficway Opposite Direction

Crash Configuration: Head-On
50 Lateral Move (Left/Right)
51 Lateral Move (Going Straight)
52 Specifics Other
53 Specifics Unknown

Crash Configuration: Forward Impact
54 This Vehicle's Frontal Area Impacts Another Vehicle
55 This Vehicle Is Impacted by Frontal Area of Another Vehicle
56 This Vehicle's Frontal Area Impacts Another Vehicle
57 This Vehicle Is Impacted by Frontal Area of Another Vehicle
58 This Vehicle's Frontal Area Impacts Another Vehicle
59 This Vehicle Is Impacted by Frontal Area of Another Vehicle
60 This Vehicle's Frontal Area Impacts Another Vehicle
61 This Vehicle Is Impacted by Frontal Area of Another Vehicle
62 Specifics Other
63 Specifics Unknown

Crash Configuration: Sideswipe/Angle
64 Lateral Move (left/Right)
65 Lateral Move (Going Straight)
66 Specifics Other
67 Specifics Unknown

Crash Category: Change Trafficway Vehicle Turning

Crash Configuration: Turn Across Path

68 Initial Opposite Directions (Left/Right)
69 Initial Opposite Directions (Going Straight)
70 Initial Same Directions (Turning Right)
71 Initial Same Directions (Going Straight)
72 Initial Same Directions (Turning Left)
73 Initial Same Directions (Going Straight)
74 Specifics Other
75 Specifics Unknown

Crash Configuration: Turn Into Path

76 Turn Into Same Direction (Turning Left)
77 Turn Into Same Direction (Going Straight)
78 Turn Into Same Direction (Turning Right)
79 Turn Into Same Direction (Going Straight)
80 Turn Into Opposite Directions (Turning Right)
81 Turn Into Opposite Directions (Going Straight)
82 Turn Into Opposite Directions (Turning Left)
83 Turn Into Opposite Directions (Going Straight)
84 Specifics Other
85 Specifics Unknown

Crash Category: Intersecting Paths (Vehicle Damage)

Configuration Straight Paths

86 Striking from the Right
87 Struck on the Right
88 Striking from the Left
89 Struck on the Left
90 Specifics Other
91 Specifics Unknown

Crash Category: Miscellaneous

Crash Configuration: Backing, Etc.

92 Backing Vehicle
93 Other Vehicle or Object
98 Other Crash Type
99 Unknown Crash Type
00 No Impact

Source: Researcher determined — inputs include police report, scene inspection, vehicle inspection, and interview.

SAS Data Set: GV
SAS Variable: DRINKING

Element Attributes:

0 No alcohol present
1 Yes - alcohol present
7 Not reported
8 [No driver present]
9 Unknown

Source: Police report

SAS Data Set: GV
SAS Variable: DRUGS

Element Attributes:

0 No other drug(s) present
1 Yes other drug(s) present
7 Not reported
8 [No driver present]
9 Unknown

Source: Police report.

Screen Name: Posted Speed Limit

SAS Data Set: GV
SAS Variable: **SPLIMIT**

Element Attributes:

	Enter posted speed limit in kmph
000	No statutory limit
999	Unknown

Range: 0-122, 999

Source: Primary sources are scene inspection or statutory law. ***Do not*** use the police report for selecting this variable's value.

Screen Name: Police Reported-Travel Speed

SAS Data Set: GV
SAS Variable: TRAVELSP

Element Attributes:

	Enter police reported travel speed
999	Unknown

Range: 0-240, 999

Source: Police report only

Screen Name: PAR Severity

SAS Data Set: OA
SAS Variable: **INJSEV**

Element Attributes:

0	O — No injury
1	C — Possible injury
2	B — Nonincapaciting injury
3	A — Incapacitating injury
4	K — Killed
5	U — Injury, severity unknown
6	Died prior to crash
9	Unknown

Source: Police report.

Screen Name: Role

SAS Data Set: OA
SAS Variable: ROLE

Element Attributes:

1	Driver
2	Passenger
9	Unknown

Source: Primary source is interviewee; secondary source is police report.

Screen Name: Time of Crash

SAS Data Set: ACCIDENT
SAS Variable: TIME

Element Attributes:

0001-2400	
9999	Unknown

Source: Police Report

Screen Name: Sex

SAS Data Set: OA
SAS Variable: SEX

Element Attributes:

1	Male
2	Female — Not reported pregnant
3	Female — pregnant - 1st trimester (1st-3rd month)
4	Female — pregnant - 2nd trimester (4th-6th month)
5	Female — pregnant - 3rd trimester (7th-9th month)
6	Female — pregnant - term unknown
9	Unknown

Source: Primary source is the interview, secondary sources include police report and official records (*e.g.* medical).

Screen Name: Date of Crash

SAS Data Set: ACCIDENT
SAS Variable: MONTH, DAYWEEK, YEAR

Element Attributes:

Month:	01-12
Weekday:	01 through 07
Year:	2002

Source: Assigned by Automated Case Selection System

DAY OF WEEK values are coded as follows:
01 Sunday 05 Thursday
02 Monday 06 Friday
03 Tuesday 07 Saturday
04 Wednesday

SAS Data Set: **GV**
SAS Variable: **RELINTER**

Element Attributes:

0	Non-interchange area and non-junction
1	Interchange area related
2	Intersection related/non-interchange
3	Driveway, alley access related/non-interchange
4	Other junction (specify) / non-interchange
5	Unknown type of junction / non interchange
9	Unknown

Source: Researcher determined — Primary source is the scene inspection, secondary sources include the police report and interviews.

MANNER OF COLLISION (SAS Label: MANCOLL)

This single place numeric value indicates the configuration of the crash based on the first harmful event, using the following codes:

0	NOT COLLISION WITH VEHICLE IN TRANSPORT
1	REAR-END
2	HEAD-ON
4	ANGLE
5	SIDESWIPE, SAME DIRECTION
6	SIDESWIPE, OPPOSITE DIRECTION
9	UNKNOWN

Screen Name: Roadway-Alignment

SAS Data Set: **GV**
SAS Variable: **ALIGNMNT**

Element Attributes:

1	Straight
2	Curve Right
3	Curve Left
9	Unknown

Source: Researcher determined--Primary source is scene inspection; secondary sources include the police report and interviews.

Appendix B: Sample EDR Report (NHTSA, 2002)

Vetronix

CDR CRASH DATA RETRIEVAL SYSTEM

CDR File Information

Vehicle Identification Number	1G1JC524317******
Investigator	
Case Number	
Investigation Date	
Crash Date	
Filename	2002-50-070-V1.CDR
Saved on	xxxxx
Collected with CDR version	Crash Data Retrieval Tool 1.331
Reported with CDR version	Crash Data Retrieval Tool 2.900
Event(s) recovered	Deployment Deployment Level

SDM Data Limitations

SDM Recorded Crash Events:
There are two types of SDM recorded crash events. The first is the Non-Deployment Event. A Non-Deployment Event is an event severe enough to ?wake up? the sensing algorithm but not severe enough to deploy the air bag(s). It contains Pre-Crash and Crash data. The SDM can store up to one Non-Deployment Event. This event may be overwritten by another Non-Deployment Event. This event will be cleared by the SDM after the ignition has been cycled 250 times.
The second type of SDM recorded crash event is the Deployment Event. It also contains Pre-Crash and Crash data. The SDM can store up to two different Deployment Events, if they occur within five seconds of one another. Deployment Events cannot be overwritten or cleared from the SDM. Once the SDM has deployed the air bag, the SDM must be replaced.
The data in the Non-Deployment Event file will be locked after a Deployment Event, if the Non-Deployment Event occurred within 5 seconds before the Deployment Event unless a Deployment Level Event occurs within 5 seconds after the Deployment Event, and then the Deployment Level Event will overwrite the Non-Deployment Event file.

SDM Data Limitations:
-SDM Recorded Vehicle Forward Velocity Change is one of the measures used to make air bag deployment decisions. SDM Recorded Vehicle Forward Velocity Change reflects the change in forward velocity that the sensing system experienced during the recorded portion of the event. SDM Recorded Vehicle Forward Velocity Change is the change in velocity during the recording time and is not the speed the vehicle was traveling before the event, and is also not the Barrier Equivalent Velocity. This data should be examined in conjunction with other available physical evidence from the vehicle and scene when assessing occupant or vehicle forward velocity change. For Deployment Events and Deployment Level Events, the SDM will record 100 milliseconds of data after deployment criteria is met and up to 50 milliseconds before deployment criteria is met. For Non-Deployment Events, the SDM will record the first 150 milliseconds of data after algorithm enable.
-SDM Recorded Vehicle Speed accuracy can be affected if the vehicle has had the tire size or the final drive axle ratio changed from the factory build specifications.
-Brake Switch Circuit Status indicates the status of the brake switch circuit.
-Pre-Crash Electronic Data Validity Check Status indicates ?Data Invalid? if the SDM receive an invalid message from the module sending the pre-crash data.
-Driver?s Belt Switch Circuit Status indicates the status of the driver?s seat belt switch circuit. If the vehicle?s electrical system is compromised during a crash, the state of the Driver?s Belt Switch Circuit may be reported other than the actual state.
-Passenger Front Air Bag Suppression Switch Circuit Status indicates the status of the suppression switch circuit.
-The Time Between Non-Deployment and Deployment Events is displayed in seconds. If the time between the two events is greater than five seconds, ?N/A? is displayed in place of the time.
-If power to the SDM is lost during a crash event, all or part of the crash record may not be recorded.
-If the vehicle is a 2000 - 2002 Chevrolet Cavalier Z24 or a Pontiac Sunfire GT, with a manual transmission (RPO MM5) and a 2.4L engine (RPO LD9), the Brake Switch Circuit Status data will be reported in the opposite state than what actually occurred, e.g. an actual brake switch status of ?ON? will be reported as ?OFF?.
SDM Data Source:
All SDM recorded data is measured, calculated, and stored internally, except for the following:
-Vehicle Speed, Engine Speed, and Percent Throttle data are transmitted once a second by the Powertrain Control Module (PCM), via the vehicle?s communication network, to the SDM.
-Brake Switch Circuit Status data is transmitted once a second by either the ABS module or the PCM, via the vehicle?s communication network, to the SDM.
-The SDM may obtain Belt Switch Circuit Status data a number of different ways, depending on the vehicle architecture. Some switches are wired directly to the SDM, while others may obtain the data from various vehicle control modules, via the vehicle?s communication network.
-The Passenger Front Air Bag Suppression Switch Circuit is wired directly to the SDM.

Vetronix

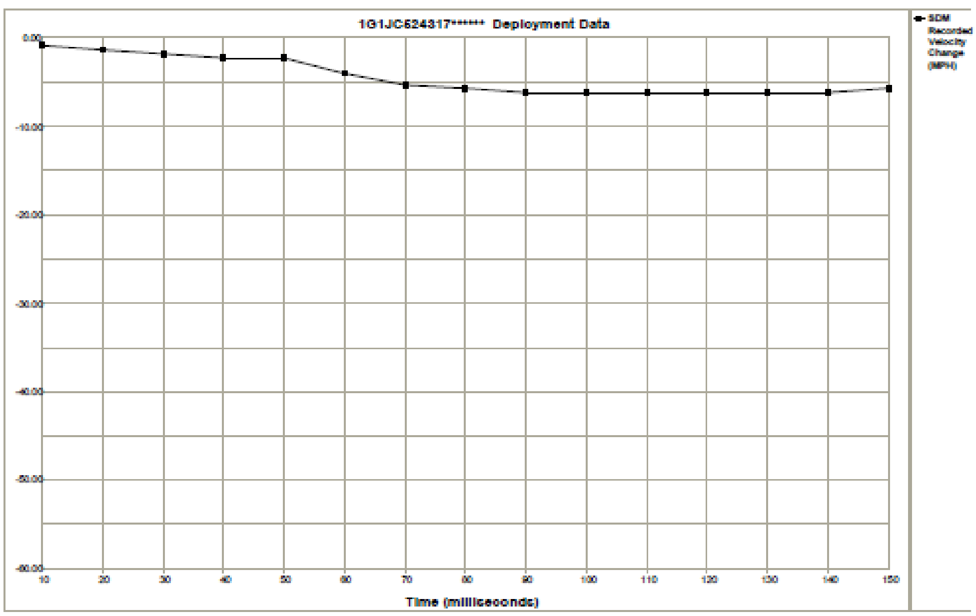
CDR CRASH DATA RETRIEVAL SYSTEM

System Status At Deployment

SIR Warning Lamp Status	OFF
Driver's Belt Switch Circuit Status	BUCKLED
Passenger Front Air Bag Suppression Switch Circuit Status	Air Bag Not Suppressed
Ignition Cycles At Deployment	3432
Ignition Cycles At Investigation	3434
Maximum SDM Recorded Velocity Change (MPH)	-6.45
Algorithm Enable to Maximum SDM Recorded Velocity Change (msec)	132.5
Time Between Non-Deployment And Deployment Events (sec)	N/A
Time From Algorithm Enable to Deployment Command Criteria Met (msec)	72.5

Seconds Before AE	Vehicle Speed (MPH)	Engine Speed (RPM)	Percent Throttle
-5	1	576	0
-4	1	576	0
-3	1	576	16
-2	1	1600	0
-1	5	640	0

Seconds Before AE	Brake Switch Circuit Status
-8	ON
-7	ON
-6	ON
-5	ON
-4	ON
-3	ON
-2	OFF
-1	OFF



Time (milliseconds)	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
SDM Recorded Velocity Change	-0.88	-1.32	-1.76	-2.19	-2.19	-3.95	-5.27	-5.70	-6.14	-6.14	-6.14	-6.14	-6.14	-6.14	-5.70

Vetronix

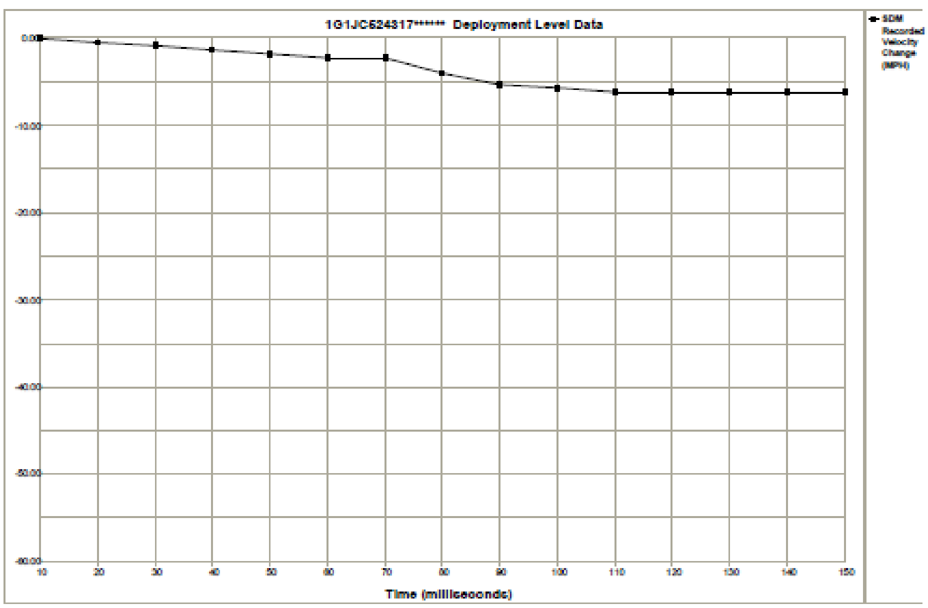
CDR CRASH DATA RETRIEVAL SYSTEM

System Status At Deployment Level

SIR Warning Lamp Status	OFF
Driver's Belt Switch Circuit Status	BUCKLED
Passenger Front Air Bag Suppression Switch Circuit Status	Air Bag Not Suppressed
Ignition Cycles At Deployment Level	3432
Ignition Cycles At Investigation	3434
Maximum SDM Recorded Velocity Change (MPH)	-6.45
Algorithm Enable to Maximum SDM Recorded Velocity Change (msec)	132.5
Time From Algorithm Enable to Deployment Command Criteria Met (msec)	72.5
Time Between Deployment And Deployment Level Events (sec)	N/A

Seconds Before AE	Vehicle Speed (MPH)	Engine Speed (RPM)	Percent Throttle
-5	1	576	0
-4	1	576	0
-3	1	576	0
-2	1	576	16
-1	5	1600	0

Seconds Before AE	Brake Switch Circuit Status
-8	ON
-7	ON
-6	ON
-5	ON
-4	ON
-3	ON
-2	OFF
-1	OFF



Time (milliseconds)	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
SDM Recorded Velocity Change	0.00	-0.44	-0.88	-1.32	-1.76	-2.19	-2.19	-3.95	-5.27	-5.70	-6.14	-6.14	-6.14	-6.14	-6.14

Vetronix CDR CRASH DATA RETRIEVAL SYSTEM

Hexadecimal Data

```
$01   08 23 00 00
$02   95 6B
$03   41 53 31 31 34 31
$04   4B 32 31 54 4C 32
$05   00
$06   22 67 40 98
$10   FE 52 FC
$11   87 03 88 FA 8F 00
$14   03 84 AB 80
$18   83 82 83 B4 FF 00
$1C   FA 32 4A FA FA FA
$1D   FA FA 32 4A FA FA
$1E   FA FA
$1F   FF 02 00 00 00
$20   A0 00 00 FF 5D F8
$21   FF BF FF FF FF FF
$22   FF FF FF FF FF FF
$23   7C 1B 01 D7 03 00
$24   01 02 03 04 05 05
$25   09 0C 0D 0E 0E 0E
$26   0E 0E 00 08 01 01
$27   01 01 00 3F 00 00
$28   2A 00 00 00 00 19
$29   09 09 09 09 00 FE
$2A   52 FF FF FF FF FF
$2B   FF FF FF 00 00 04
$2C   00 00 22 03
$2D   35 37 27 1D
$30   A0 00 00 FF 5D F8
$31   FF BF FF FF FF FF
$32   FF FF FF FF FF FF
$33   7C 22 03 03 0E 0D
$34   02 03 04 05 05 09
$35   0C 0D 0E 0E 0E 0E
$36   0E 02 1B 01 D7 08
$37   01 01 01 01 00 3F
$38   00 00 00 2A 00 00
$39   00 0A 19 09 09 09
$3A   00 FE 52 FF 32 00
$3B   00 04 00
$3C   1D 35 37 27
$40   FF FF FF FF FF FF
$41   FF FF FF FF FF FF
$42   FF FF FF FF FF FF
$43   FF
```

REFERENCES

Evans, L. (2004). *Traffic safety*. Bloomfield Hills, MI: Science Serving Society.

National Highway Traffic Safety Administration. (2008a). *Traffic safety facts 2006: Overview*. Washington, DC: National Center for Statistics and Analysis, DOT HS 810 809.

National Highway Traffic Safety Administration. (2005). *Analysis of speeding-related fatal motor vehicle traffic crashes.* U.S Department of Transportation, DOT HS 809839.

National Highway Traffic Safety Administration. (2007). *Traffic safety facts 2006: Speeding.* Washington, DC: National Center for Statistics and Analysis, DOT HS 810 814.

National Highway Traffic Safety Administration. (2008b). *Analysis of event data recorder data for vehicle safety improvement.* U.S Department of Transportation, DOT HS 810 935.

German, A. (n.d.). Event Data Recorders: A new resource for traffic safety research?. Retrieved June 29, 2008, from http://www.nhtsa.com/ staticfiles /DOT/NHTSA/ NRD/ Articles/EDR/PDF/Research/Event_Data_Recorders _New_Resource_ for_Traffic_Safety_Research.pdf.

Gabler, H.C., Hampton, C.E. & Hinch, J. (2004). *Crash severity: A comparison of event data recorder measurements with accident reconstruction estimates*. Society of Automotive Engineers, Technical Paper Series. No. 2004-01-1194.

Chidester, A., Hinch, J. & Roston, A.T. (2001). *Real world experience with event data recorders*. National Highway Traffic Safety Administration, Paper Number 247.

Chidester, A., Hinch, J., Mercer, T.C. & Schultz, K.S. (1999). *Recording automotive crash event data*. International Symposium on Transportation Recorders. Arlington, VA.

Evans, L. (1991). *Traffic safety and the driver*. Van Nostrad Reinhold, NY: Science Serving Society.

Kloeden, C., McLean, J., Moore, V. & Ponte, G. (1997). *Traveling speed and the risk of involvement*. University of Adelaide, Road Accident Research Unit, Adelaide.

Garber NJ & Gadiraju R. (1988). *Speed variance and its influence on accidents*. Washington, DC: AAA Foundation for Traffic Safety.

Federal Highway Administration. (1998). *Synthesis of safety research related to speed and speed management.* Retrieved July 15, 2008, from http://www.fhwa.dot.gov/tfhrc/safety/pubs/speed/spdtoc.htm.

Fitzpatrick, K., Carlson, P., Brewer, M., Wooldridge, M. & Miaou, S. (2003). Design speed, operating speed, and posted speed practices. *National Cooperative Highway Research Program, Transportation Research Board, Report 504.*

Solomon, D. (1964). *Accidents on main rural highways related to speed, driver and vehicle.* Washington, DC: Federal Highway Administration.

Cirillo, J. (1968). Interstate system accident research study II. *U.S. Bureau of Public Roads. Interim Report II, Vol. 35, No. 3,* 71-75.

Joksch, H. C. (1993). Velocity change and fatality risk in a crash-A rule of thumb. *Accident Analysis and Prevention, Vol. 25, No. 1.*

Research Triangle Institute. (1970). *Speed and accidents: Volume I.* Washington, DC: US Department of Transportation.

O'Day, J. & Flora, J. (1982). *Alternative measures of restraint system effectiveness: Interaction with crash severity factors*. Society of Automotive Engineers, Technical Paper Series. No. 820798.

Bowie, N.N. & Walz, M. (1994). *Data analysis of the speed-related crash issue*. Auto and Traffic Safety.

Hauer, E. (1971). Accidents, overtaking and speed control. *Accident Analysis and Prevention, Vol. 3, No. 1.*

West, L. & Dunn, J. (1971). *Accidents, speed deviation and speed limits*. Traffic Engineering, 52-55.

Harkey, D., Robertson, H. & Davis, S. (1990). Assessment of current speed zoning criteria: Record, No. 1281. *Transportation Research Board, Washington, D.C.*

Ogle, J. (2005). *Quantitative assessment of driver speeding behavior using instrumented vehicles.* Georgia Institute of Technology.

National Highway Traffic Safety Administration. (2002). *National automotive sampling system crashworthiness database: 2002-2003 coding and editing manual.* Washington, D.C: National Center for Statistics and Analysis.

National Highway Traffic Safety Administration. (n.d.). *National center for statistics and analysis: Available data*. Retrieved July 20, 2008, from ftp://ftp.nhtsa.dot.gov/NASS/.

National Highway Traffic Safety Administration. (1995). *National accident sampling system crashworthiness database: Data collection coding and editing manual.* Washington, D.C: National Center for Statistics and Analysis.

National Highway Traffic Safety Administration. (1998). *National automotive sampling system crashworthiness database: Analytical user's manual*. Washington, D.C: National Center for Statistics and Analysis.